Hormonal modulation of developmental plasticity in an epigenetic robot

John Lones

Department of Computer Science
University of Herfordshire

This dissertation is submitted in partial
fulfilment of the requirements of the degree of
Doctor of Philosophy

June 2016

Acknowledgements

I would like to thank my Supervisors Lola Cañamero and Matthew Lewis in regards[1] to their support and guidance during During My time and the University of Hertfordshire.

In Addition I would also like to thank past and present members of the EECAiA Lab and Academic support staff

Finally I would also like thank friends and family for their support particularly my other-half Zozi, and my parents Freda and Nick

---

[1] sorry :)

Abstract

In autonomous robotics, there is still a trend to develop and tune controllers with highly explicit goals and environments in mind. However, this tuning means that these robotic models often lack the developmental and behavioral flexibility seen in biological organisms. The lack of flexibility in these controllers leaves the robot vulnerable to changes in environmental condition. Whereby any environmental change may lead to the behaviors of the robots becoming unsuitable or even dangerous.

In this manuscript we look at a potential biologically plausible mechanism which may be used in robotic controllers in order to allow them to adapt to different environments. This mechanism consists of a hormone driven epigenetic mechanism which regulates a robot's internal environment in relation to its current environmental conditions.

As we will show in our early chapters, this epigenetic mechanism allows an autonomous robot to rapidly adapt to a range of different environmental conditions. This adaption is achieved without the need for any explicit knowledge of the environment. Allowing a single architecture to adapt to a range of challenges and develop unique behaviors.

In later chapters however, we find that this mechanism not only allows for regulation of short term behavior, but also long development. Here we show how this system permits a robot to develop in a way that is suitable for its current environment. Further during this developmental process we notice similarities to infant development, along with acquisition of unplanned skills and abilities. The unplanned developments appears to leads to the emergence of unplanned potential cognitive abilities such as object permanence, which we assess using a range of different real world tests.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Hormone Regulation of developmental plasticity

The physical and behavioral characteristics of animals have been established as being the result of the evolution process of natural selection [35, 141]. Yet over the last 50 years we are becoming increasing aware that variations in characteristics can occur in relation to the organism's internal and external environmental conditions (see [76, 124, 105, 42, 146, 101, 148]). This would suggest that the developmental road map is not concrete, with environmental exposures and sensory experiences that occur during the animal's life having the potential to influence the phenotype [63]. The ability for an organism's phenotype to be affected by exposures may potentially provide a critical adaptation mechanism by allowing the organism to adapt to unfamiliar environments and environmental changes, highlighting the inherent behavioral flexibility and plasticity found in animals.

Unlike the behavioral flexibility found in animals however, in autonomous robotics, there is still a trend to develop and tune controllers with highly explicit goals and environments in mind (see [135, 70] for an overview). This tuning can be very direct such as pre-determining the weighting of environmental cues, or more subtle through the use of mechanisms such as reward feedback, fitness functions and activity functions [70, 79] However, even slight changes in the environment can lead to significant and often unpredictable changes in the robot's behavior ([125, 15, 133, 86].

While environmental changes of course also affect the trajectory of biological organisms behaviors (see [29, 32, 149]), unless the changes are extreme, the organisms will survive. In contrast for robotic controllers which utilize pre-programed/determined adaptation mechanisms even minor changes can lead to behaviors that are not only unsuitable but may render the robot inoperable or even dangerous [125, 15, 133, 86, 140, 70, 146]. In order for autonomous robots to overcome this limitations they will likely need to possess an open-ended development system in which the robot's abilities can continuously extended and be refined as it interacts with its environment. However, the question is how can such a system be implemented into an autonomous robot?

Like with many recent innovations in Artificial intelligence, one answer to an open-ended development system may lie in nature. As previously stated animals have been shown to posses a mechanism(s) that allow environmental exposures to alter their phenotype. Known as phenotype plasticity, this phenomena describes the extent of said animal is able to change characteristics ranging from behavior to morphology to physiology in response to an environmental cue [42, 146]. One way in which environmental cues are able to modulate development is through the endocrine system [91, 42, 146, 84, 38]. For example stressful experiences lead to variations in corticosteroids levels, if the variation or "imbalance" is large enough, it can lead to long lasting cognitive and behavioral change (see [50, 91, 92]). A potential problem though is that the exact mechanism(s) that facilitate the said adaptation are not entirely clear. That being said, one potential which has recently gained a large amount of attention is that of varying endocrinology based epigenetic mechanisms.

In recent years, a large body of work has emerged describing ways in which hormone can influence development through varying "epigenetics mechanism". These mechanisms consists of histone modification and DNA methylation which are shown to regulate gene expression throughout the body and subsequently alter an organism's behavior. While their is still some debate over the exact trigger of these mechanisms [32], evidence has shown hormones are correlated to changes in gene expression and exposure to different environmental factors (e.g., [33, 106, 32]. It is therefor possible that this style of "epigenetic" mechanism may provide a way in which environmental cues through hormones can influence the development and developmental plasticity of an animal. A mechanism which regulates developmental plasticity may also have significant advantages for an autonomous robot. By allowing environmental factors and experiences to shape development, significant adaptive gains and further autonomy may be achievable. Specifically this type of mechanism may reduce the need to design controllers with limited scope which are only able to operate in very specific conditions.

However, before we discuss the potential for a hormone based epigenetic mechanism as a tool for adaptation, it would be first useful to define exactly what we mean by an epigenetic mechanism as the term can take different meaning.

1.2 Defining epigenetics: The origins of epigenetics

In the academic community there are divisions over the exact definition of what constitutes epigenetics, with multiple independent origins readily identifiable ([56, 59]). It should be stressed that due to the focus of this manuscript we barely brush the surface of the current state of epigenetic research.

The term "epigenetic" itself is borrowed from ancient Greek and in its most literal form means the “over/above/upon which brings into being”. In today’s academic community anything termed “genetic” is likely to be associated with the life sciences. Yet as if to typify the issues with defining epigenetics, the term epigenetic likely arose in the earth science field of geology where it refers to the emergence rock formations. Specifically in the geology context epigenetics is the study of mineral deposits which are younger than the encasing rock i.e how the surroundings rocks leads to the origin/emergence or “birth” of the deposit(e.g., [45]).

It was not until later that Waddington [144, 145] who is often considered the "grandfather" of modern epigenetics used the term to describe the study of epigenesis, the general development and morphogenesis of organisms. Waddington [145] and others such as [55] considered mechanism that allowed identical cells to develop along different trajectories to be epigenetic, leading to the metaphor of an "epigenetic landscape" (see figure 1.1). Although the underlying mechanism was not yet known, similar to the earth science approach to epigenetic, [145] believed that the cell’s environment (i.e exposures to temperature or different chemicals) was responsible for controlling the developmental trajectory [59]. For instance, Hadorn *et al.* [55] identified how identical imaginal discs in *Drosophila* (fruit fly) larvae had the potential to develop into different appendages, such as antennae, wings or legs depending upon the disc’s surroundings. This research represented some of the earliest work to demonstrated that it is not only genetic but also epigenetic mechanisms which regulated gene expression, cellular differentiation and the overall development processes.

More recently Waddington’s approach to epigenetics and development has been defined as “the branch of biology which studies the causal interaction between genes and their products which bring the phenotype into being” [65, 59].

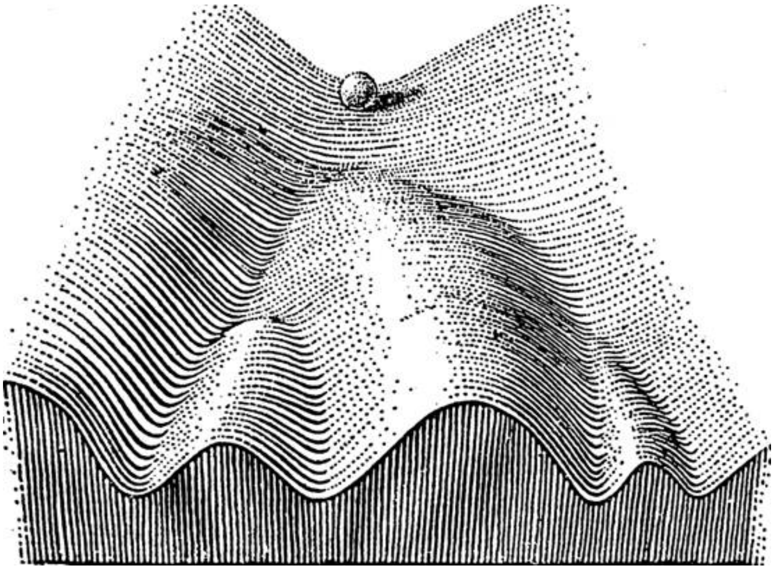


Fig. 1.1 Waddington's epigenetic landscape.
Waddington's epigenetic landscape is a metaphor used to described how environmental factors can influence gene regulation. According to Waddington gene regulation can be compared to a marble rolling down a hill, in the sense that the path it roles down to reach its eventual resting point is influenced by groves in the hill. In a similar manner he argued the genes final resting point (the genes expression) was influenced by the groves (environmental factors) it was exposed down its pathway (development), with Waddington coining the term chereode to describe this process.

1.2.1 A new direction: Epigenetics and disease susceptibility

While Waddington's approach to epigenetic's resulted in many discoveries, the term fell out of usage with developmental biologists and as whole remained relatively unused, it was not until 1975 when both Holliday and Pugh [60] and Riggs [120] revisiting Waddington's proposals, began to look into the role that DNA methylases had on gene regulation within *eukaryotic* (cells with a nucleus) organisms.

While not every aspect of their original proposed model proved accurate (see Haig [56], R [118]), later findings by Bird [14] brought found more evidence pointing to methylases as providing a role in gene regulation. This cumulated in 1996 with Russo, Martienssen and Riggs [121] redefining epigenetics in the narrower context of the study of heritable changes in gene expression that cannot be explained by genetic change. However, it was still a few more years before Nan *et al.* [104] bought definitive evidence of methylases role in gene expression. This discovery marked a resurgence of interest in epigenetic mechanics ([65, 59]. While this new form of epigenetics did still branch into general areas of biological development (e.g [49]), it became increasingly associated with disease and cancer research [65]. Of particular interest were the roles of epigenetic mark and mutation on disease susceptibility (see [66]). As developmental biologists had stopped using the term epigenetics or epigenetic terminology in their work, Russo *et al.* [121] definition soon became adopted as the primary meaning [65, 56, 59].

1.2.2 Epigenetics and a return to developmental plasticity

More recently epigenetic research has once again expanded back into development, seeking to both provide a framework and understand how the expression of genes is influenced by environmental factors and the resulting effect on behavior [107, 31, 43, 32, 116], cognition [43, 37, 36, 73], personality [67, 9], mental health [139] and morphology [100]. This form of epigenetics has been defined in broader manner and includes a collection of diverse phenomena with definitions such as epigenetic being “any non-genetic mechanism that influences the phenotype” [30]. This form of epigenetics is sometimes known as *“behavioral epigenetics”* [31, 28]

A critical note to make is that while these “developmental epigenetic changes” can be hereditary, a majority of studies focus on the effects that these changes have on the current generation. For example [107, 31, 43, 131, 73] focus of either early life experience during development i.e., pre/post-natal periods or later lifestyle choices. This is in contrast with more traditional epigenetic based studies such as those discussed in section 1.2.1 were heritability is a critical component. This has recently become controversial [98, 147], with arguments rising if this approach is “epigenetic” in nature. Due to confusion and controversy, it would not be of a total surprise if this form of epigenetics “re-branded” itself in the near future.

For the moment, under this developmental approach to epigenetics in certain cases for certain organisms, to have any influence, the environmental stimuli that trigger the subsequent epigenetic mechanism needs to be present during what is called the critical development period [119, 43]. Depending upon the species, this critical period can incorporate the early prenatal stages and can finish before birth or continue into the postnatal stage. Examples of this phenomenon have been documented in insects [100], reptiles [123] and mammals [107]

In cases such as Shine and Downes [123] and Miyakawa *et al.* [100], as the epigenetic changes have a significant morphological or behavioral aspect to them, they provide accessible examples of how the immediate environment can dictate the development along certain paths.

In other examples, particularly those focusing on the development of cognition in mammals, the effects of epigenetic changes tend to be less measurable. Taking one of the most well studied examples, the effects of stress in pre- and postnatal periods (e.g [107, 131] it has been found that exposure to stressors during these critical periods can lead to avoidance behavior in later life.

However, epigenetic change is not just limited to early development. Examples of epigenetic mechanism working in later life have been discovered. For example Human monozygotic (identical) twins have been shown to be epigenetically identical during youth, yet in later life significant long after the critical period has passed epigenetic difference have been found to emerge [49]. Another potential example is stress disorders. Recent studies has found some evidence of an underlying epigenetic process behind the exposure to stress and behavioral and mental issues [20].

Studies in mice further demonstrate that epigenetic changes can an do occur in later life as a response to an identifiable trigger. For example exposure to stress or nutrition variations in adult mice is shown to result in different epigenetic variations (see, [72])

However, again while it is possible to measure and observe the effects of exposure to different environmental stimuli on development, there is still some uncertainty over the exact mechanics(see [32]). Yet while their are uncertainties over the exact mechanism, a promising lead (which we will discuss in greater detail in section 2.4.2) consists of hormones providing signals for the epigenetic mechanism.

Before we go into any more detail on the potential role of hormone in epigenetic mechanism, their are few more aspects of epigenetic that needs be discussed. Firstly under this new developmental approach to epigenetics there are at least two variations in the mechanisms, context dependent epigenetics and germline epigenetics.

1.2.3 Context dependent epigenetic change

Context dependent epigenetic changes are temporal (and usually molecular) changes that occur as a result of exposure to certain environmental factors. These changes can at a later date be reversed or modified by either removal of, or changes in the factor. The ability for a change to occur and be reversed or modified is not uniform, with the type of factor, duration and timing all likely to play a role. While certain changes such as those discussed in Miyakawa *et al.* [100] will rapidly occur and reverse, other changes (see [31, 32, 20] can persist long or potentially indefinitely after the factor has been removed. Factors which are commonly linked to these types of epigenetic change range from maternal care [107, 27] to diet [20] to exposure to heavy metals [61] (see Crews [32] for a more complete overview).

At the current time there is uncertainty if these changes are truly trans-generational, in the sense they are directly inherited. Long standing opinion has held that these changes would not be passed down through the generations. However, recent studies have provided some evidence that context dependent epigenetic changes can leave epigenetic markers which are heritable and can lead to behaviors being directly passed down through generations [33]. Additional there is also the potential, for certain changes to be passed down cross-generational (non hereditary) through channels such as maternal care or imitation (see [27, 57, 74]). For instance, a mother with a withdrawn phenotype from exposure to a factor, may be less attentive to her young, leading to them also developing a withdrawn phenotype without exposure to the original factor. In this sense the parent becomes the environment factor. Of course with cross generation transmission, other factors will also influence and mediate any transmission [74].

1.2.4 Germline epigenetic changes

Germline epigenetic changes consist of changes which are “truly” trans-generation in which change through mechanisms such as DNA methylation or histone modification become embedded into, and passed down through the germline [32]. As truly trans-generational changes, these are passed down and appear without the need for exposure to the original causation factor. These types of changes are associated with prenatal development variations such as sex, height or eye color as well as issues relating to disease susceptibility [44].

1.2.5 The epigenetics in artificial life: Epigenetic robots

Epigenetics itself is not a foreign concept in artificial life and robotics community. Like the study of epigenetics in the natural world, epigenetics within the robotic community also has divisions. These divisions can lead to confusion over what exactly is meant by epigenetics robots and mechanisms. Unfortunately unlike with epigenetics of the natural world which has clear conceptual divisions, in our community epigenetics has recently seem to become synonymous and interchangeable with the term developmental robotics .

If we look through some of relevant definitions this lack of focus is hardly surprising due to their relatively broad nature,

- Epigenetic Robotics, "is a new discipline at the frontier of developmental psychology, neural-, and engineering sciences whose goal is to model the development of cognition in natural and artificial systems" Berthouze and Metta [13]
- Epigenetics "encompasses a broad spectrum of issues, investigating the acquisition of motor skills and the role played by morphological development" Lungarella *et al.* [83].
- "Epigenetic robotics is one of the new cognitive modeling approaches to modeling autonomous mental development" Cangelosi and Riga [24].
- "Epigenetic robotics research builds models of the psychological development of children" Prince *et al.* [117].

In addition, to the broad nature of the definitions, in many cases models under the epigenetic robotics banner, possess no systems which could be considered a biologically plausible epigenetic mechanism. Rather these models tend to be of a relatively high level conceptual design with little or no focus on genes.

Piagetian epigenetics and epigenetic robots

So why the choice of the term epigenetic robotics? As previously discussed, the lack of a genetic underpinning makes the choice of "epigenetics" seem strange for this subsection of development robotics. However, the reasoning and justification for the adoptions of the term "epigenetic robotics" can be linked to the psychologist Jean Piaget. Piaget, who himself had previously borrowed the epigenetic terminology from Waddington [144], particularly the concept of the epigenetic landscape, to explain his theory on human behavioral development, "genetic epistemology" ([114] see [64]). In its simplest form, Piaget's theory of development suggests that a child's cognition emerges as a result of both genetics and its interactions with its environment, for instance the concept of sensory motor development.

Since epigenetic robotic models aim to facilitate an open ended ongoing development process, which like Piaget theory, allows environmental interactions to shape development, at the 2001 First International Workshop on Epigenetic Robotics (Now ICDL-Epirob) it was argued that the usage of the term *epigenetic robotics* was justifiable [1, 25].

However, this does lead to a few concerns,

- Firstly, Piagets argument that his theory of development fell with following Waddington's[144] definitions of epigenetic, are not widely accepted. Many leading publication into both epigenetics and the history of epigenetic do not acknowledge Piaget work as being epigenetic in nature (for example see [59] or [56]). Those such as Jablonka (see [65] or [64]) that do refer to Piagetian ideas, still have their own research highly grounded in genetics.

- Secondly, while Piaget borrowed terminology, he did so only to describe observation in development. As we demonstrate ourselves in chapters 6 and 7, at least some of his observations may emerge due to a biologically plausible epigenetic mechanism, and therefore could be a epigenetic phenomena. As epigenetic robotics often try to implement new high level mechanisms which replicate Piagetian or other developmental theories, they critically remove the presence of "genes", or to put simply remove the possibility for a underpinning epigenetic mechanism. .

1.3 A new developmental epigenetic robotic model

The recent research in epigenetic has pointed to the existence of a mechanism possessed by animals which may permit environmental stimuli to modify the phenotype and regulate developmental plasticity. Research into the roles of epigenetic mechanism on developmental plasticity is still in its infancy with most studies focusing on understanding the mechanisms and the implications on the phenotype and behavior in a general setting (see section 1.2.2). This is not to suggest that investigations on the influences of environmental factors on development is new (see[146]) and has not been explored in robot models (see [83, 25]). In fact a large body of work has already been established in the robotic community to investigate ways in which environmental cues and developmental plasticity and be utilized, including models of behavior [26], perception [83], personality [130] and self-organization [39]. However, in our model we take a relatively different approach as will become clear in subsequent chapters. However, to briefly summarize,

- Firstly, we try to limit explicit motor control, instead we allow behaviors to emerge as a natural phenomena of development.
- Secondly we ground development in the environment and body. This is achieved by giving the robot survival related needs, i.e., the robot learns a certain behavior because it finds it beneficial given its environment and it morphology.
- Third we not only propose both a novel mechanism to introducing developmental plasticity into a model, using a hormone model, but this mechanism is also biologically plausible.

The use of a epigenetic approach to regulate developmental plasticity has so far only been investigated in a limited number of studies, particularly in cases of adaption (See [142, 94, 46, 88]). However, those that have framed have shown promising potential.

In this work, we begin to look at the potential for a biologically plausible epigenetic like mechanism to be used as a tool to create adaptive behavior in an autonomous robot by regulating development. Like with the epigenetic mechanisms found in animals (See section 1.2.2), the epigenetic mechanism proposed here is designed to allow internal and external stimulation to influence the development and phenotype of the robot.

It should be stressed that the research and model proposed here is of an abstract take of how roles epigenetic mechanisms may play in the regulation of development. We do not, nor claim to model a highly detailed epigenetic mechanism i.e., we do not explicitly model DNA methylation of genes. Rather, we use abstract models to represent this process. This decision was made as in early research we found more detailed models being too computationally heavy in our simple autonomous robots. In terms of behavior and development which is the focus of this manuscript, early simulation showed limited variation between our abstract model presented here and more detailed models.

As the focus here is on the effect of these stimuli on the current generation, our work bear more similarities to context dependent epigenetic mechanism as discussed in section 1.2.3. Whereby environmental factors are able to modulated and change the development of robot through the epigenetic mechanism effect the phenotype of the current generation.

However, in order for such a mechanism to work, the robot must firstly be able operate autonomously in an open environment to become exposed to different environmental stimuli which trigger our epigenetic like mechanism. Additionally and perhaps more importantly the epigenetic mechanism proposed here exists to modulate both behavior and the robot's phenotype, it therefore needs to be integrated into a base autonomous architecture which it is able to modulate.

In order to achieve the desired automation we take a well documented biologically inspired approach based on an analogy to homeostatic control. These systems as described in the following chapter 2 requires the robot to maintain a safe internal environment. Using a homeostatic controller, as the robot's internal environment moves away from ideal state, intrinsic motivations emerge in order to attempt to return to the ideal state. It is these motivations which are used in order to get the wanted automation.

1.4 Hypothesis

The Hypothesis of this research is that a biological plausible epigenetic mechanism is able to increase the embeddedness and coupling of an autonomous robot with its current environment by exploiting internal and external cues. Specifically we propose that this mechanism will, through hormone concentrations, be able to use these cues in order to regulate developmental plasticity of a robot, leading the emergence of unplanned behaviors and motivations which are suitable for its given environment. Further we suggest that such an epigenetic mechanism can provide a foundation for embodiment for which further advanced systems, such as cognitive development, can be integrated into. As these systems are integrated and incorporated into the epigenetic architecture we suggest that these systems can be modulated both directly and indirectly.

- Directly in sense that the epigenetic mechanism can control and modulated the systems
- Indirectly in the sense that as the epigenetic mechanism regulates the robot's behavior and interactions.

In order to verify this Hypothesis we must first explore the potential adaptive benefits of the epigenetic system and it ability to facilitate adaption appropriately to environmental cue, in which,

1. The proposed architecture should be able to facilitate the generation and emergence of motivations and behaviors in order to allow independent automation. Meaning that the robot's adaption is neither constrained by the designer's level of knowledge of both the robot and its future environments.
2. The proposed epigenetic mechanism should be able to demonstrate the generation of motivations and behaviors by modulating and adjusting internal systems of the architecture in relation to environmental conditions cues and the robot's own internal conditions. In order for this to occur, the internal hormonal concentrations which drive and signal the mechanism.
3. To claim adaptability, the architecture must demonstrate that not only can it modulate its behaviors and motivations in relation to environmental cues, but that the modulation improves performance/survivability . Here performance of the robot is measured by comparing different architectures in a range of environments while considering the following.
    - The ability to manage its internal needs

- The ability to exploit ecological niches
- The ability to exploit it own morphological advantages
- The ability to recognize and respond to temporal changes and dynamic events.

4. The proposed architecture must be flexible, with the same architecture able to adapt to a range of different environmental conditions and challenges. A task driven robot may show adaptive capabilities in a specific environment, however taken out of it sterile environment they are often rendered inoperable (see Krichmar [71]).

   In addition if we have ambition of utilizing the architecture in a dynamic real word environment the robot must also demonstrate the ability to....

   - Generate and develop unplanned behaviors and motivations which are suitable for its current environment. For instance while robots can be programmed with predetermined behaviors, if these behaviors are fixed or coarse grained then they may become unsuitable if the environment changes. The same can also be said about motivations and goals, as depending upon the environmental conditions the proprieties or suitabilities of different motivations may no longer be relevant.
   - Be able to respond appropriately and adapt to the unknown. Regardless of how much knowledge of the world a robot is endowed with, it is likely to come across new or novel environments, situations and interactions, it is therefore important that a robot has some for of mechanism to allow it to respond and adapt to these in an appropriate manner.

Secondly once we have established that the robotic model and epigenetic mechanism permits environmental cues to be used as adaptive signals, we then look into the potential roles of an epigenetic mechanism on cognitive development, by considering the following

1. Whether the proposed architecture is able to regulate learning in relation to both the internal and external environment in an appropriate manner.
2. Is the robot able to prioritize which elements of the environment it should attempt to learn about first?

1.5 Thesis Structure

The thesis is organized in the following manner

- Chapter 2 describes the first implementation of our epigenetic model. This includes both the description of the robot itself along with the hormone system and the epigenetic mechanism.
- Chapter 3 introduces the first experiments which were conducted in order to examine the ability of the epigenetic mechanism to regulate the development and adaptation of an autonomous robot in a range of different environments.
- Chapter 4 considers the limitations of the epigenetic mechanism found in the previous chapter, where in order to operate the robot relies on the designer to provide the robot with the affordances of different objects. We suggest that a curiosity based learning mechanism can overcome these limitation. This chapter looks at how the current hormone based neuro-modulation model can be extended to allow the robot to identify and interact with novelty thought the emergence of curiosity based motivations.
- Chapter 5 continues the previous chapter by introducing the learning component of the curiosity driven learning mechanism, a Artificial Neural Network which we have called an emergent neural network or ENN. The ENN allows the robot to learn from the interactions and behaviors which are generated by the neuro-modulation.
- Chapter 6 investigates how the combination of the ENN and neuro-modulatory systems allows a robot to adapt to a range of environments and situations including human-robot interaction without the need of any pre-requisite knowledge or parameters
- Chapter 7 builds upon previous chapters were we noticed interesting developmental phenomena in the robot. Here we notice how the robot appears to go through unplanned developmental stages that bears some similarity to the concept of sensory motor development. Specifically in this chapter we investigate both the potential cognitive abilities of the robots using well know tests, as well as looking at what role the robot's environmental conditions play in the emerge of this “cognition”.
- Chapter 8 provides a summary of the research reported and both a look into additional and potential future work .

1.6 Publication List

The following is a list of paper produced while working on this PhD:

1. Lones, J. and Cañamero, L. "Epigenetic adaptation through hormone modulation in autonomous robots." 2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL). IEEE, 2013.
2. Lones, J. Lewis, M. and Cañamero, L.. "Epigenetic adaptation in action selection environments with temporal dynamics." Advances in Artificial Life, ECAL 2013: Proceedings of the Twelfth European Conference on the Synthesis and Simulation of Living Systems. 2013.
3. Lones, J. Lewis, M. and Cañamero, L.. "Hormonal modulation of interaction between autonomous agents." 4th International Conference on Development and Learning and on Epigenetic Robotics. IEEE, 2014.
4. Lones, J. Lewis, M. and Cañamero, L.. "Hormonal modulation of development and behaviour permits a robot to adapt to novel interactions," , ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems, Vol. 14. New York, 184–191. 2014
5. Lones, J. Lewis, M. and Cañamero, L. "From Sensorimotor Experiences to Cognitive Development: Investigating the Influence of Experiential Diversity on the Development of an Epigenetic Robot." Frontiers in Robotics and AI 3 (2016): 44.

Additional and future publication can also be found at

https://scholar.google.co.uk/citations?user=lMilqIYAAAAJ&hl=en

Chapter 2

A biologically-plausible epigenetic mechanism

In this chapter we will begin our research into a potential adaptive epigenetic mechanism by first considering how to implement such a system in an autonomous robot. As we previously discussed in section 1.2.2, the exacts mechanism behind this epigenetic phenomena are not fully understood. However, research has found and a linkage between hormones and change in gene expression (e.g., [47, 32] see section 2.4.2). Suggesting that changes in hormones levels could well provide a trigger for epigenetic change. Before we go further however two key points need to be made,

- Firstly, the epigenetic process that occurs in biological creatures is likely a complex phenomenon which involves a range of different processes and chemical modulators. In this implementation, we have simplified a potential plausible mechanism with the focus on a few key components.

- Secondly, as we will shortly discuss, while the focus in this thesis is on the roles of hormones in the epigenetic process, we do not wish to suggest that these are the only causes or triggers, merely that they could be one potential .

With that being said, since hormones have so far been shown to be related to the activation of epigenetic mechanisms in biological organisms, here we utilize artificial hormones to trigger our epigenetic mechanism. This epigenetic mechanism results in changes in gene-like expression, effecting the robot's hormone receptors, whereby high concentration of a hormone leads to the up-regulation of the receptors (i.e., an increased sensitivity of receptors) and lower concentrations cause down-regulation (i.e., a lowered number of receptors). In this sense the epigenetic mechanism discussed bear similarities to finding of [107, 149]

An epigenetic change in regulation of hormone receptors will change the robot's sensitivity to the different hormones circulating around its body. Therefore, in cases of a highly up-regulated receptor, only a fraction of the hormone compound would be needed to elicit the same response as in a robot with unmodified receptors. This provides the robot with an ability to automatically regulate its internal systems to the current environment, potential creating a more adaptive agent. However, before we go into greater about the epigenetic system we will first discuss the robotic model which this system is integrated into, a homeostatic control system. As we previously eluded to in chapter 1 in order for the epigenetic mechanism to work, it must be integrated into an autonomous architecture.

2.1 Homeostatic controllers and robots

At the heart of our robotic model lies a survival based homeostatic system. This system consists of 3 variables that the robot must maintain within a predetermined boundary of $0 < x < 100$ esstentially creating a physiological space [93] or viability zone [5]. Transgression of any boundary will result in the "death" of the robot. Here "death" means that the robot will essentially shut down performing no new actions or undergoing no new internal changes with no potential to re-start. In order to manage these needs and survive the robot requires a "allostasis" like mechanism (see [134, 90]). in which its behavior, desires and motivation adjusts in response to its current homeostatic condition in order to try to return to balance. Here we implement a form of allostasis by having motivations generated and modulated in relation to the robot's homeostatic conditions. Originating from the works of Taylor [136] who himself based his concept on Ashby's concept of ultrastability (see [41]), using homeostasis as an intrinsic mechanism to generate motivations has previously been demonstrated as a successful adaptive mechanism in a range of studies (E.g., [3, 17, 40, 6, 18]. Perhaps one of the earliest and best known implementations in a robot was achieved in Kismet [17]. Kismet, which consisted of an arousal based human robot interaction model using a disembodied robotic head, had facial expression which were regulated by the type of stimulation it was exposed to and its current internal state. For example over-stimulation or stimulation when the robot has low desire for stimulation led to the robot displaying a "disgusted" expression.

In later works such as [41] and [6] the use of homeostatic controls was expanded to include autonomous robots. A crucial contribution of these models is not the automation per se, but the inclusion of the concept of viability and survival. As Di Paolo [41] states “survival is the mother of all variables”. By using survival related variables, in order to survive these robots are both required and have the ability to adjust their behavior to suite both their environments and internal environment. In the case of Avila-García *et al.* [6] the homeostatic controller was demonstrated to be able to appropriately manage multiple unrelated needs.

In this model we use these survival related internal variables in a similar way as Di Paolo [41] and Avila-García *et al.* [6]. Here it is the level of the survival related variables that directly or indirectly drive everything the robot does, be it movement, interaction and in later chapters even learning (see chapters 6 and 7). In some cases the relationship between behaviors and these variables can easily be established and identified. For example a behavior that directly recover one of the variables. However in other cases as the robot learns and adapts to its environment the link is less clear. For example in chapter 6 we discuss the robot attacking a perceived threat. In the short term this behavior led to a reduction in the robot overall homeostasis, however in the long term, by driving off the threat the robot enjoys overall higher homeostatic levels.

2.2 Our homeostatic system

The homeostatic system consists of homeostatically-controlled variables representing different internal needs that the robot must maintain within a pre-determined range between [0,100] in order to survive. Each of the survival-related homeostatic variables has a lethal boundary which if transgressed results in the robot’s death. In the case of energy and health, the lethal boundary is set at the bottom end of the range of permissible values, in the case of temperature the lethal boundary is at the upper end of the range. These variables create a viability zone (see [5]) within which survival is guaranteed, and a physiological space that permits us to measure quantitatively the “quality” of the robot’s survival and adaptation in different ways [6].

In this model, the survival related variables that the robot needs to maintain within the viability zone are energy, health and internal temperature. These three variables were chosen as they represent in our opinion, 3 of most relevant and critical survival needs that an autonomous robot would need to manage. Where possible we have attempted to link these needs to the robot’s physical shell.

- The first homeostatic variable is energy (e), which is a measure of the robot's available battery charge. During these experiments, as we were unable to implement a method to directly charge the battery during a live run, we instead created a simulated battery that the robot could sense internally. With this simulated battery, the robot had a maximum charge of 100 units and would lose one unit of charge per second. In order to recharge this battery, the robot would need to find and consume "energy resources" (see figure 2.1), which consisted of pink balls located throughout its environment. Consumption of these resources would occur automatically if the robot was within a 2 cm range and facing the ball.

- The second homeostatic resource is health (h), which decreases in a semi-dynamic manner in relation to physical contact, which is detected using the IR sensors. The health variable is used to represent the physical health or condition of the robot, i.e., chassis damage. Obviously for practicality and monetary issues, it is not reasonable to truly damage the robot during our experiments hence this use of IR sensors to detect contact. Much like energy, this homeostatic need can be recovered through the consumption of certain resources, in this case blue colored balls. For both the health and energy variables, if the robot's homeostatic level drops below 0, the robot "dies" (stops operating).

$$Health\Delta = \sum_{i=0}^{15} \begin{cases} \frac{Ir_i}{IrMax} \times persistence, & \text{if } Ir_i > (0.8 \times IrMax). \\ 0, & \text{otherwise} \end{cases} \tag{2.1}$$

  Where Ir_i is the value from infrared sensor i, $IrMax$ is the maximum potential value of the ir sensors (the value when an object is detected at point blank range). *Persitance* is the duration the object is detected for, in terms of action loops.

- The final homeostatic variable is internal temperature (t). Internal temperature is related to both the current speed of the robot (with increased speed leading to the generation of additional heat) and the external environmental temperature or "climate", as can be seen in equation 2.2.

$$TemperatureGain = \frac{|sp|}{10} * Climate \tag{2.2}$$

where $|sp|$ is the current absolute value of speed of the wheels (measured in rotations per action loop) and 10 is a predetermined constant to regulate temperature gain in relation to movement.

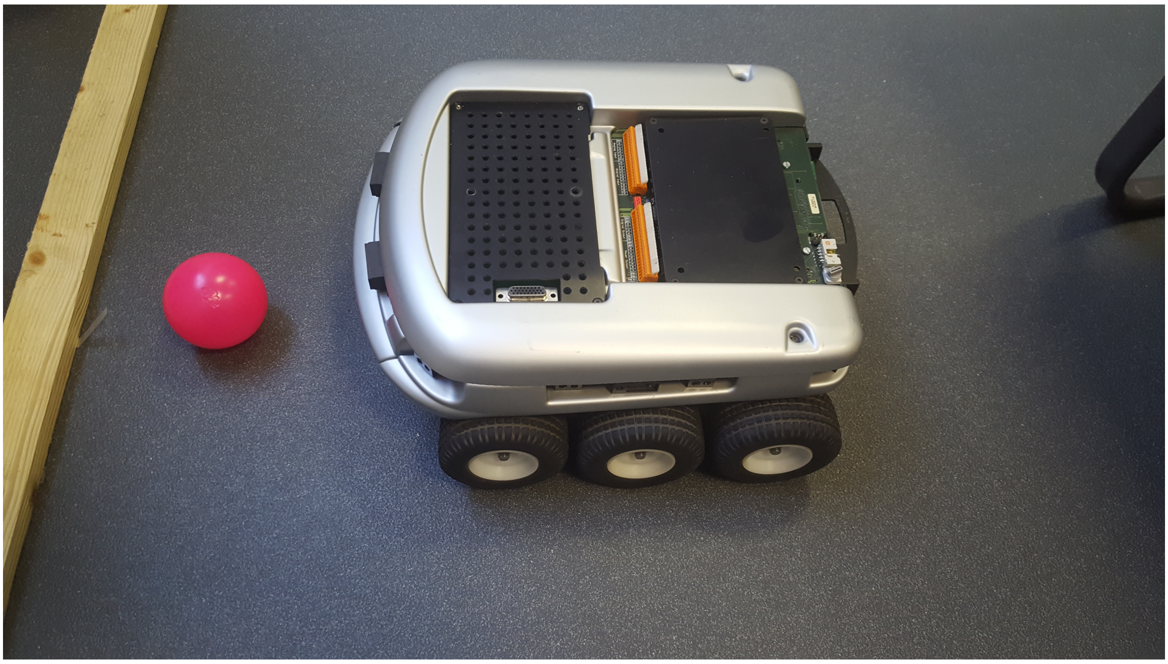


Fig. 2.1 The Koala II robot “consuming” a energy resource

The robot’s body temperature is set to dissipate at a constant rate of 5 percent of total current temperature per action loop. Dissipation is the only direct method available to the robot to reduce its body temperature, meaning that in order to cool down, the robot must either reduce or suppress movement. At the same time, the only method for the robot to gain heat is to increase its movement speed. Unlike with the “health” and “energy” variables, in this case robot tried to maintain a low temperature, with death occurring if its temperature reaches a value of 100.

Table 2.1 The homeostatic variables

H.Var	Ideal Value	Limit	Cause of Defcit
Energy	100	0	Constant loss of 1 unit per second.
Health	100	0	Physical contact (see equation 2.1)
Temperature	0	100	Movement (see equation 2.2)

2.3 Homeostatic control through hormonal modulation

As previously mentioned, homeostatic variables have proven to be a useful tool to generate behaviors and motivations in autonomous robots in simple environments. However, purely motivation-driven models have been shown to perform poorly in more complex or dynamic environments [7] where the robot is required to balance multiple needs.

A potential solution to this problem is again found in biological systems. In animals, homeostatic motivations do not come directly from deficits. Rather, the motivation and ability for the animal to sense the need is signaled by different chemical messengers, particularly hormones. Hormones derived from homeostatic deficits (e.g., ghrelin in the case of hunger) are shown to be behind the onset of motivation by providing a signal of the deficit [85] and the motivational value of environmental cues [52].

Like with homeostatic variables, hormone based mechanisms in which hormones are secreted in relation to changes in a robot's internal (e.g., the presence of a homeostatic deficit), or external environment (e.g., the detection of a predator) have been proven to give rise to short-term, rapid adaptation to a range of different environmental dynamics. Once secreted, these hormones, acting upon selected aspects of the architecture such as in [7, 70]) modify the behavior and perception of the robot by "rewiring" different neural or physiological functions. Similarly to natural hormones in biological systems, the artificial hormones decay over time, allowing the robot to revert to its default behavior once the need or challenge has passed.

However, while these architectures offer increased behavioral flexibility and adaptability, they are not without limitations. So far most hormonal models have only addressed simple reactive behaviors such as avoidance or attraction and are constrained to simplistic —even though dynamic— environments. In addition, these models also require the designer to have a good understanding of environmental conditions in order to configure different aspects of the hormonal system such as the rates of secretion and decay.

While the hormonal-/neuro-modulation model that we propose shortly is somewhat unusual in the field of developmental robotics, our system does share some common ground with more widespread models such as [58, 108]. In these models, as in our own, research is often devoted to the investigation of mechanisms that support, and have a potentially crucial role in, the emergence of cognition. These systems have focused for the most part in modeling novelty and curiosity, as well as different mechanisms underlying the emergence of sensory-

Table 2.2 The different hormones used in our architecture.

Hormone	Group	Secretion trigger	Promotes motivation
E1	eH	Energy defcits	Hunger
H1	eH	Health defcits	Repair
T1	eH	High internal temperature	Hyperthermia
D1	nH	Visual cues and eH concentration	Dominant behavior

motor skills [**?**]. This work, however, also presents a few major difference with respect to these approaches.

- Firstly, we have no explicit expectation of an outcome or end state. While many hormonal based developmental systems (see [**?**] for an overview) try to mimic a specific developmental phenomenon, we are interested in how this system modulates the behavior of the robot in a more open-ended, exploratory investigation.
- Secondly, rather than trying to explicitly model theoretical constructs, as is often the case in developmental robotics, in our work we use hormone to create a system that allows these “constructs” to emerge unplanned as a natural aspect of the development of the robot.

In this system, hormones secreted from glands simulated in the model, modulate different sensors and actuators of the robot. As we will show later, these hormones are secreted in relation to both changes in environmental conditions as well as changes in the robot homeostasis. The amount of secretion in relation to changes is directly related to the size of the stimuli. In our neuro-modulatory model, we have implemented two different types of hormones, which are classified as either endocrine hormones (eH) or neuro-hormones (nH) in Table 2.2.

2.3.1 Endocrine hormones

Drawing on biological systems, our eH-like hormones are implemented with the primary purpose to help maintain internal homeostasis. The eH group is comprised of three hormones, each one associated with one of the three homeostatic variables. These hormones are then secreted in relation to the current homeostatic deficits as shown in equation 2.3:

$$eHSecreation_v = \psi_v \times Deficit_v \tag{2.3}$$

where ψ_v is the activity level of hormone gland v, a static value to regulate secretion.

2.3.2 Hormone receptors

Once secreted, the hormones modulate the robot is varying ways (see section 2.4) until they decay (see formula 2.4). Hormone receptors located in the action selection architecture are able to detect the different concentration of these two types of hormones, which is determined by the total sum of the active secretions. These receptors are then able to generate internal drives which are later used to calculate the robot's motivations. Although the robot has three homeostatic needs, and therefore three drives, we only need to directly calculate the intensity of "hunger" (tendency to consume energy) and "repair" (tendency to recover health). "Hyperthermia" (tendency to reduce temperature), which is satisfied by reduced or no movement, instead suppresses other drives as can be seen in equation 2.5.

$$\Delta eHConcentration_v = eHSecretion_v + (eHConcentration_v * .95) \tag{2.4}$$

Where $eHConcentration_v$ is the current concentration.

$$Drive_i = \frac{Sens_v eH_v}{Sens_t eH_t} \tag{2.5}$$

Where $Sens_v$ is receptors sensitivity to eH_v the hormone concentration where i $\neq$ t i.e., V is energy or health, but not temperature.$Sens_t$ and eH_t refer to the sensitivity to and concentration of the T1 hormone.

2.3.3 Neurohormones

The second group or hormones, neurohormones (nH), during this stage of the model contains only one hormone, D1. This hormone regulates what can be described as "dominant" or potentially "aggressive" behavior, and can be thought of as being somewhat similar to Testosterone in mammals. This "dominant" behavior is achieved by having the hormone suppresses environmental cues that are associated with negative stimuli (stimuli associated with a move away from homeostasis such as physical contact or overheating) as well as increase the robot's speed (shown later in formula 2.9).

For example, as a side-effect of the concentration level of D1 and its effects on perception of stimuli, a robot with a high D1 level that detects a desired resource will move directly towards it at a high speed, pushing aside any obstacles and disregarding the potential of damage from collisions. In contrast, a robot in the same situation but with a low D1 level would instead move around obstacles to reach the desired location.

As we have previously discussed in section 2.2, increased speed (and/or a hotter climate) raises the temperature of the robot and therefore leads to a motivation to reduce speed as a way to regain homeostatic balance. However, with a high *nH* concentration, the robot is able to greatly suppress its tendency to reduce its temperature, allowing increased movement speed in times when it is needed.

This type of system shares some similarities with [7, 8], where a hormone suppressed perceptual feedback from the robot's bumpers (touch sensors), causing the robot to push competitors away as a side effect. The main difference in the system that we present here is that the *nH* hormones are used to suppress any negative stimuli. In addition, in our model the strength of suppression of negative stimuli is correlated to the concentration of hormones present in the system, as in [21], rather than with internal deficits, providing a simple form of "affective memory".

The secretion of D1 is determined by total internal *drive*, as given by the formula:

$$nHSecreation_{D1} = \sum_{v}(motivation_{v}) \tag{2.6}$$

where $motivation_i$ is the current motivation to replenish a survival related variable which is shown later in formula 2.8 in section 2.4.2.

Once secreted like the eH hormones, the concentration of D1 is detected by a specific hormone receptor ($sens_{D1}$) to determine the current modulating effect of hormone which we will refer to as dominance.

$$dominance = nH_{d1} \times Sens_{d1} \tag{2.7}$$

where nH_{D1} is the concentration of the neurohormone D1

2.4 The robot architecture

With the hormone system able to signal the internal needs of the robot, it now needs an architecture which it is able to modulate. This architecture which we have designed to incorporate the hormones modulation is a simple sensory driven model. Unlike more common hormone based architecture sch as [22, 8, 71], in our case the hormone concentrations are not used to select or activate a pre-defined states or behavior. This robot model has no pre-defined states or behavior to switch or select from.

Rather, behaviors emerge due to the modulation of actuators and sensors in relation to internal and external stimuli and signals, as we will discuss shortly. However, before the robot becomes able to modulate its internal systems, it must first discover its surroundings. This is achieved through the use of both the infra-red (IR) sensors and a camera (web-cam) as discussed in the following section.

2.4.1 The robot

Experiments were conducted using the medium-sized wheeled robot Koala II (see figure 2.2). The Koala II robot, which is equipped with 16 infrared sensors located around its body. These IR sensors are used to detect both the distance and presence of any nearby objects, as well as in-lieu of physical touch sensors to detect contact. This can be seen in part of the previous equation 2.1. In addition the robot was equipped with a Logitec™C270 hd webcam located on top of the robot and pointing towards its front to permit the use of vision, and a laptop running the Ubuntu operating system, that we placed on top of the robot for it to be able to move around freely rather than being limited by the length of the serial cable.

In addition to these senses, the robot also possesses limited color-based vision using OpenCV and the web-cam. This permits the robot to detect the presence, location and size of the different resources needed to satisfy homeostatic needs. Since all resources in the environment are of the same size, the robot is able to judge both the distance and the accessibility (i.e., if a resource is partially hidden) through the size of the stimuli.

Figure 2.3 shows and example of the robot's vision: the detection of an environmental cue, where i is the type of cue and d is its position. As shown at the top of the image, cues are grouped into one of five directions, from far left to far right of the robot's field of vision. The term "valence" refers to the intrinsic attractiveness or averseness of a detected cue, which is determined as a function of both the size of the cue and the current internal needs of the

Fig. 2.2 The Koala II robot (without attached laptop) used during the experiment.

robot, and has an effect on the behavior of the robot. For example, a perceived high positive valence will encourage the robot to move towards the object.

2.4.2 The sensory-driven model

With the surroundings detected, our sensory-driven model is now able to modulate behavior in an appropriate manner given the current environmental conditions. The first step of this modulation takes into account the current internal state and the environmental conditions in order to allow for the emergence of "motivational" states. Here we use the term "motivation" to describe, not a predetermined hard-coded state which activates under certain conditions, as usually done in the Action Selection literature (see [6]), but a hypothetical construct to describe the robot's behavior. These emerged "motivations" are dependent upon both the internal drive of robot—e.g., a need to recover energy—and the current environmental cues—e.g., the detection of an energy source. As "motivations" emerge due to the interaction of different drives and environmental conditions, in theory it is possible for a wide range of different "motivations" to occur. In practice, in the environments that we have used, three primary motivations appeared regularly and can be readily identified, as shown in table 2.4.

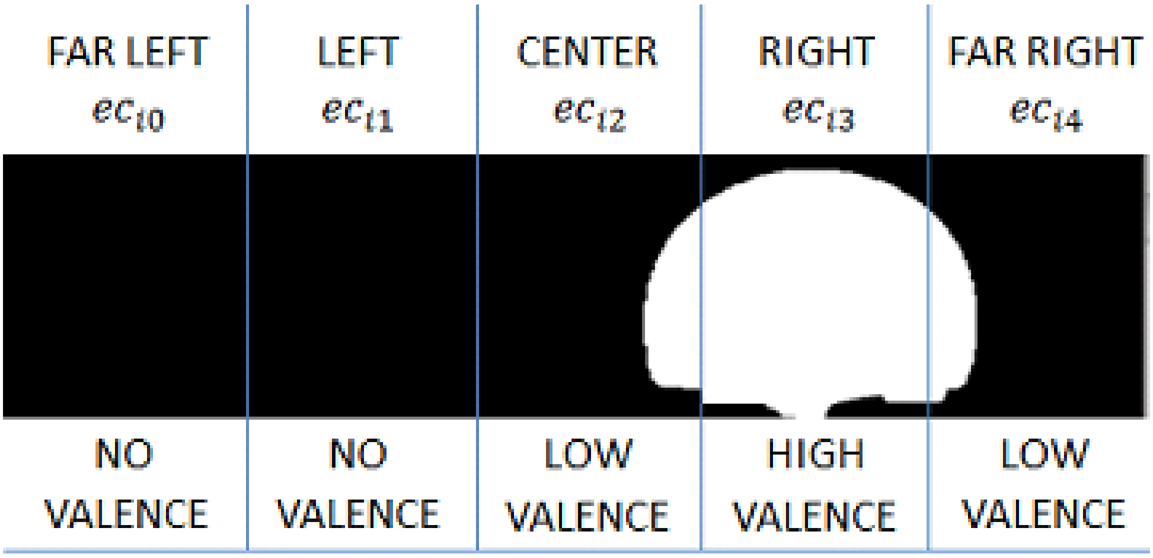


Fig. 2.3 Example of how the robot sees
Using the openCV contour and color detection the robot can detect different environmental cues $(ec_{i,d})$. i is the type of cue (i.e., in this example an energy resource) and d is the position. Cues are grouped into one of five directions, from far left to far right of the robot's field of vision. The term "valence" refers to the perceived attractiveness or aversiveness of a detected cue.

Once again although the robot has three homeostatic needs, we only need to directly calculate the motivations to recover energy and replenish health. The motivation to reduce temperature emerges naturally due to the suppression that occurs in formula 2.5. Here motivations are calculated in a similar manner to [93] as shown in the following

$$Motivation_{v,d} = drive_v + (drive_v \times environmentalCue_{v,d}) \tag{2.8}$$

where $Motivation_i$ is the motivation to recover a homeostatic need i, $enviromentalCue_{i,d}$ is the detection of an environmental cue of the environment that effects the $homeostaticVariable_i$ either positive or negative , with d the position of that cue.

As can be seen in the equation 2.8, the robot's "motivation" to move is driven by both its internal state and external stimuli. Using this algorithm, in cases when the robot's internal needs are satisfied, it will engage in an avoidance-like behavior and try to maintain a distance between itself and the objects perceived, stopping if it finds an empty[1] area.

---

[1] Due to the Koala's sensor range, an "empty area" this roughly translates to an area of 25 square centimeters.

Table 2.3 The different environmental cues the robot was programed to detect, and the effect they had on the robot's homeostasis. A solid object refers to any object detected by an individual I-R. Therefore the robot can detect and be affected by multiple solid objects at any given times

Objecy	Effect on variable		
	Health	Energy	Temperature
Health resource (blue ball)	+10	0	0
Energy resouce (pink ball)	0	+10	0
Solid object	-5	0	0

Table 2.4 Motivations regularly observed in our robot.

Motivation	Drive	Ext Stimuli	Suppressed by	Promoted behaviours
Hunger	Recover Energy	Energy resource	High Temp/D1	Energy seeking
Repair and avoidance	Recover Health	Health resource	High Temp/D1	Repairing and avoiding objects
Hyperthermia	Reduce Temp	External Climate	Low energy/health/D1	Reduced speed

In contrast, when the robot needs to replenish health or energy but no external stimuli are available to do so (i.e., no resources are detected), the robot will engage in an "exploratory" behavior. This exploration is an emergent behavior achieved when the robot moves forward and avoids objects whenever they are detected.

The second step in the sensory driven model determines the robot's behavior generate behavior (motor action) given the current motivational state. Rather than having pre-determined behaviors or "activities", different behaviors emerge in this architectures due to the interaction of fine-grained elements. For example, rather than having an explicit "pushing" behavior, the robot may push an object when its "desire" to move forward outweighs its "desire" to avoid the object.

As behaviors are not explicitly modeled, they do not have predetermined costs or gains. Rather, the costs or gains of any behavior can be calculated as the sum of physiological changes that occur during the execution of the action. Therefore, the cost of the said "push" behavior will be dependent on a range of factors such as speed of the robot (the temperature cost) as well as weight and size of the object to be pushed (the health cost see equation 2.1).

The model thus calculates the motion of the robot rather than specific behaviors. Motion is calculated independently for each wheel and the overall movement (its speed and direction) results from the simultaneous movement of both wheels. The speed of the wheels, and therefore the observable behavior is calculated by:

$$\begin{aligned} WheelSpeed_k = \sum_v (motivation_{v,d} \times DirectionScaling_{k,d} \times dominance) \\ D = L, R(direction) \\ DirectScaling_L = (-2, -1, 0, 1, 2) \\ DirectScaling_R = (2, 1, 0, -1, -2) \end{aligned} \tag{2.9}$$

where $wheelSpeed_k$ is the speed of the left or right wheel,$DirectScaling_{k,d}$ is the scaling of the motivation depending upon the direction it is associated with.

Hormones as epigenetic signals

Now that we have an basic autonomous architecture that is able to explore it environment, it is now possible to consider how to implement the desired epigenetic like mechanism. As previously stated while it is possible to measure and observe the effects of exposure to different environmental stimuli on development in animals, there is still some uncertainty over the exact mechanisms. Recent studies have however shown evidence that hormones may provide a signal for epigenetics [47].

Since hormonal variations and imbalances have long been known as a key modulator of short-term value-laden survival and social behaviors [68], it is not unreasonable to hypothesize that organism may have evolved to use these imbalances for long-term modulation. Recent studies have helped lend credibility to this hypothesis, as changes in gene expression in prenatal and postnatal development have strong correlation with hormone exposure [32, 47]. Other studies such as [48] have also shown that life experiences that are related to hormonal imbalances – for instance exposure to prolonged stress and cortisol levels – also have strong correlation with changes in gene expression. Therefore, while as we have said the exact mechanism that causes changes in the gene expression is not fully understood, it does appear that hormones play a role in gene expression.

2.4.3 A biologically plausible epigenetic mechanism

Since hormones appear to be related to the activation of certain epigenetic mechanisms which influence developmental plasticity in biological organism (see section 1.2.2), here we utilize the artificial hormones to trigger our epigenetic mechanism. As previously discussed at the start of this chapter, the epigenetic mechanism implemented here uses hormones to influence the expression of the robot's hormone receptors similar to the findings of [27, 32, 149]. Here

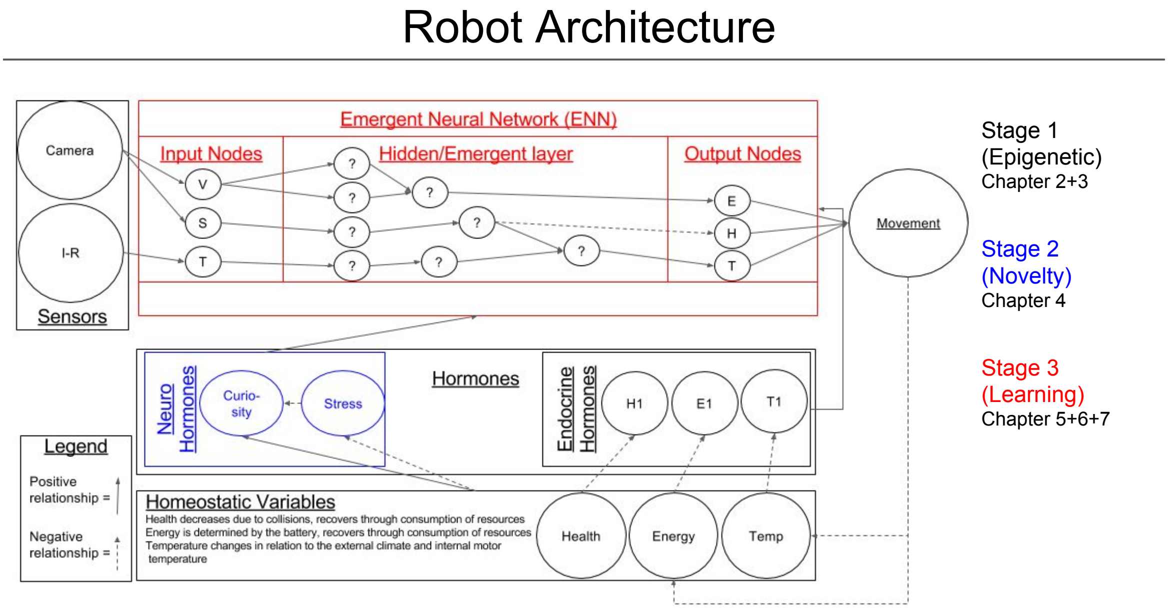


Fig. 2.4 The Robotic model

we model this epigenetic process by having high concentration of a hormone leads to the up-regulation of the receptors (i.e., an increased sensitivity receptors) and lower concentrations cause down-regulation (i.e., a lowered sensitivity receptors). As previously discussed in section 1.3, we only implement a high level model of this mechanism.

In this model we have implemented the epigenetic mechanism by having hormone receptor sensitivity ($Sens_i$) shown in formula 2.5 change as a result of exposure to its hormone concentration in the following manner

$$\Delta Sens_i = \frac{hormoneConcentration_i}{\sigma} \tag{2.10}$$

where $Sens_i$ is the hormone receptor which is effected by the concentration of its associated hormone $hormoneConcentration_i$. Resulting in the up or down regulation of the receptors. Here σ is a predetermined constant value that regulates the speed of the epigenetic process. While in biological organisms the hormone concentration would trigger a chain reaction of different chemicals to facilitate the epigenetic process, for simplicity sake here we have modeled it as a direct causation.

2.5 Conclusion

In this chapter we have introduced and discussed the basic principles of our proposed epigenetic system. This system is based on a hormone modulated, survival driven autonomous robot. In order to survive this robot is required to seek out and consume different recourses. In order to help it achieve this goal a plausible hormone driven epigenetic like mechanism which regulates the robot's hormone receptors. We propose that this system will allows a autonomous robot to adapt to range of different environments without the need for fine tuning or specific parameters. In chapter 3, we will investigate if this mechanism brings about the desired adaptation by testing this architecture in a range of different environments. In addition in order to provide a comparison, we will also test two other robotic models.

Chapter 3

The potential for adaptation using an epigenetic mechanism: Testing an autonomous epigenetic robot in a range of environments

In the previous chapter 2 we discussed how a potential hormone driven epigenetic mechanism could be implemented into a autonomous robotic model. As we have suggested in chapter 1, in biological organism such epigenetic mechanisms may provide animals with a form of adaptation through behavioral modulation. In this chapter we examine the potential for our previously described epigenetic mechanism.

The focus of this chapter and the research question centers on if the proposed epigenetic mechanism can take advantages of environmental niches by modulating the robot's internal environment in an appropriate manner given its current External conditions. By exploiting environmental niches in this way we suggest that a robot will be able to appropriately regulate development in relation to its condition. In particular we will be focusing on two aspects on the development effect of the proposed epigenetic mechanism.

- Firstly, we will be looking at how a range of different environmental conditions interplay with the epigenetic mechanisms and the effect that this has on a robot's behaviors.

Table 3.1 The three Robot models tested

	Sensory-Driven Robot	Hormone-modulated Robot	Epigenetic Robot
Sensory-Driven Layer	X	X	X
Hormone Layer		X	X
Epigenetic Layer			X

- Secondly, we will compare the adaptation and behavior of a robot possessing the epigenetic mechanism with other robotic models, allowing us to highlight the effects of the mechanism.

3.1 The robotic models used and the environments

In order to test the potential adaptive capabilities of our proposed epigenetic mechanism previously described in chapter 2, a robot possessing the mechanism was tested against two other robotic models (see section 3.1.3), a neuro-modulatory (hormone modulate) architecture (see section 3.1.3 for the model overview) and a sensory driven architecture (see section 3.1.2 for the model overview). The two later robotic architectures, the hormone and sensory driven, essentially consist of striped down version of the full epigenetic model as discussed in their relevant section (see table 3.1).

This approach, using the three models, allows us to identify and compare the explicit benefits and potential adaptive capabilities of the epigenetic mechanism. Here we can both demonstrate that the adaptive capabilities arise due to the epigenetic mechanism rather than another aspect of the model. This is important as previously discussed, both sensory driven models and hormones models have already been proven to provide adaptive benefits (see section 2.1).

In order to measure the robot's adaptability we place them in a range of different environments and measures both their survivability and overall wellbeing and viability (see section 3.2). As discussed earlier in section 2.2 the robots are required to maintain 3 survival related homoeostatic variables. It is the ability maintain these variables that will provide us one of the main performance measures.

During this first stage of experiments the different robotic models were tested in a total of six environments, which while each consisted of a three source problem, each possessed unique challenges due to environmental variation as discussed in detail in section 3.2 and surmised below.

- Environment 1: The first environment (see section 3.2.2) took place in a basic environment with easily accessible resources.
- Environment 2: The second environment (see section 3.2.3) took place in an environment with light movable object blocking the robot's path.
- Environment 3: The third environment (see section 3.2.4) took place in an environment where the resource where no longer static and would move around the environment.
- Environment 4: The forth environment (see section 3.2.5) took place in a environment with a temporal climate which was "hot" during the day and cold at "night".
- Environment 5: The fifth environment (see section 3.2.6) took place in an environment with uneven resource distribution.
- Environment 6: The sixth environment (see section 3.2.7) took place in an environment with temporal resources which were only present and accessible at certain times.

3.1.1 The three robotic models tested during the first experiments

In this section we will set out the three different robotics models used during the first six experiments. These architectures consist of different stages of the full epigenetic model, which means while they have many similarities they also have some key differences as will be discussed in the following section.

3.1.2 The Homeostatic sensory-driven robotic model

The sensory-driven architecture represents the most stripped version of the three robotic models tested here with both the epigenetic and hormone components removed. In this form the architecture has some conceptual similarities to a Braitenberg vehicle [16] where the robot's actuators are modulated in relation to sensory stimulation which is gathered from both the camera and the I-R (see section 2.4.2). However, unlike a traditional Braitenberg vehicle, the strength of any sensory stimulation is modulated by the robot's current internal environment (i.e the current levels of its internal variables) rather than the actuators being directly modulated by the stimuli, this approach bears some similarities to investigations such as [51].

The modulation of behavior via sensory stimulation still occurs through the emergence of drive and motivations as discussed in the epigenetic architecture (see section 2.4.2) however due to the lack of hormones, drives are instead implemented directly using the level of the internal variables as can be seen in equation 3.1.

$$Drive_v = \frac{Deficit_v}{Deficit_t} \tag{3.1}$$

where $drive_v$ is the desire to replenish $internalVariable_v$ depending upon the variables current $deficit_v$ and given the current internal temperature $deficit_t$. Although the robot has three homeostatic needs, and therefore three drives, like with the other models (see equation 2.5 we only need to directly calculate the intensity of "hunger" (tendency to consume energy) and "repair" (tendency to recover health). "Hyperthermia"(tendency to reduce temperature), which is satisfied by reduced or no movement, instead suppresses other drives as can be seen in the equation.

The lack of hormones in this architecture represents the main difference in comparison to both other models. The implication for this is that internal drives are based purely on the robot's current internal state, rather than being modulated by different hormones concentrations which had previously been secreted in relation to the robot's recent interactions and experiences.

Additionally this architecture also lacks the epigenetic mechanism meaning the robot's internal environment will not be modulated by exposure to different environmental factors.

With the drive calculated, the motivation to move remains the same as the epigenetic robot as shown previously in equation 3.1.

3.1.3 The neuro-modulatory model

The neuro-modulatory architecture consist of the sensory-driven architecture integrated with the artificial hormones described. The combination of these two components leads to the creation of a neuro-modulation type architecture, in which the robot's motivations and behavior are modulated by the different concentration of hormones. Neuro-modulation systems such as this have previously been implemented into a range of different robotic models such as [8, 71] whereby they have been shown to lead to efficient short term adaptation to a range of different environments. As with our Sensory driven model described in section 3.1.2, the neuro-modulatory model here is also unlike many neuro-modulation systems such as [6], due to the lack pre-defined states or behaviors.

The implementation of this model is essentially identically to the epigenetic architecture described in the chapter 1, the only difference being this models lacks the epigenetic mechanism from section 2.4.3. Without the epigenetic mechanism, hormone receptors sensitivity (see equation 2.10) would remain constant, making the explicit modeling of these receptors redundant. Therefore in this model hormone receptors have been removed meaning the previous equation 2.7 which calculated internal drives being changed to equation 3.2.

$$Drive_v = \frac{eH_v}{eH_t} \tag{3.2}$$

The epigenetic robotic model

Finally the epigenetic architecture represents the full model integrating both sensory-driven and hormone architectures along with the epigenetic mechanism as described in chapter 2.

3.2 Experiments

The six environmental conditions for which the different experiments took place in a 2m×2m arena with plywood borders which can be seen in figure 3.1. Each of our first four experiments was comprised of ten runs for each robot architecture, with a duration of 10 minutes per run, corresponding to 10,000 steps (action selection loops) in total or 16 steps per second using the microprocessor of the laptop. The last two experiments involved and 15 runs of the same duration for each architecture. The addition of 5 extra runs for last two experiments is due to increased environmental variation under these two conditions as discussed in greater detail in Sections 3.2.6 and 3.2.7. During each experiment, the robot had access to at least one of each of the homeostatic resources, with the exact number and environmental challenges changing significantly across experimental sets. If at any point in time one of the robots failed to maintain its homeostatic needs, it would be considers to have “died”, the run would end prematurely.

3.2.1 Performance measures: Defining “comfort” and “Wellbeing”

In order to analyze the performance of our three architectures, we use a range of both qualitative and quantitative metrics. Following [7], quantitative metrics are based on the

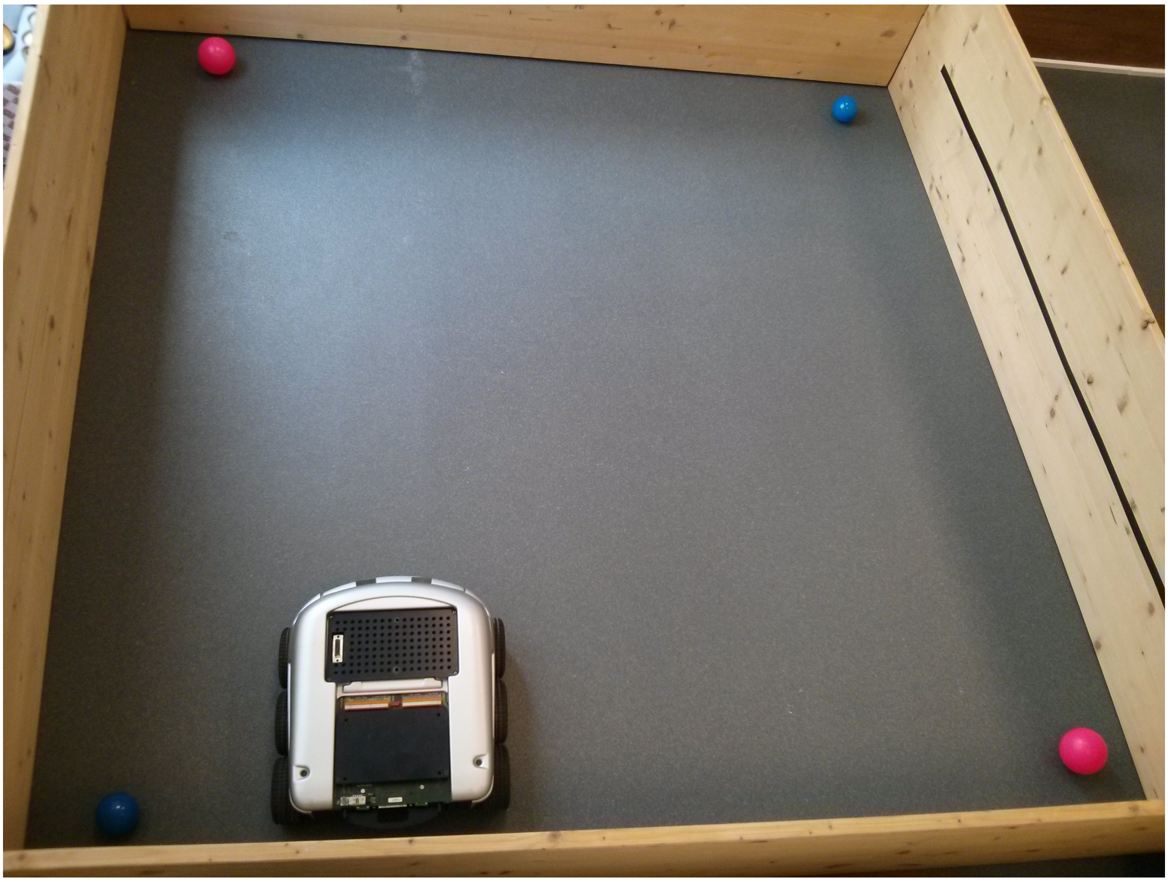


Fig. 3.1 The basic, unmodified arena used in our experiments.

notions of viability and stability of the "internal milieu" [5], and provided different ways to assess "how well" the robot maintained homeostatic balance as measured by changes of the homeostatic variables in the physiological space that they define. Here in our first experiments we use a notion of "comfort" and later (in chapters 6 & 7) we use a more flexible measure called "wellbeing".

Firstly we will discuss comfort which is used in this chapter. Comfort which has also been used in other homeostatic robot architectures (see [7]), provided a measure of the average homeostatic deficits at any time. This was calculated on a scale from 0 to 1, with a comfort level close to 1 indicating homeostatic variables near their ideal levels and a comfort level near to 0 indicating large homeostatic deficits. For example a robot with all 3 homeostatic variables at 20% of their ideal value, would have a low comfort of 0.2. This low comfort indicating the robot was close to death.

However, while we found comfort to be an adequate measure in the early experiments, later we found it to be inadequate. This inadequacy is due to an emerging phenomena in the robot resembling tolerance. Depending upon the environment, the robot is able to develop a "tolerance" to its homeostatic needs. For example, in an environment with plentifully energy, robots will develop to be less sensitive to their energy need as they adapt to being able to easily and rapidly replenish this need. The benefit for the robot is by suppressing needs and forgoing easy opportunities they focus on harder or less common needs i.e., hypothetically in a normal environment the drive to replenish a need may occur when a value drops to 70% of its ideal value. However, in an energy rich environment where the robots' sensitive to energy is diminished the energy deficit may need to reach 50% before the drive kicks in. Giving the robot more time to attend other needs.

These tolerances mean that comfort in certain cases is no longer an accurate measure of performance. For example in a environment rich with resources a robot may tolerate larger homeostatic deficits, and therefore have a lower comfort level even though its behavior may be optimal for its environment. This is particularly relevant in later chapters i.e., chapter4 onwards, where the robot has more drives and motivations then simply recover and maintenances of its survival needs

For this reason during this manuscript we often use a general measure of performance called "Wellbeing". Essentially this our "The Designer's" and "Experiment's" current view on the robot's state and situation. In order to asses the robot's wellbeing we use a range of different internal readings from the robot. For example, not only would we consider the level of homeostatic variables but also current motivations. In order to track and measures these different variables all internal values including but not limited to variables/drives/motivations/hormone & sensory readings for each action loop are printed out and saved as a cvs files. These files are then used to help analysis the robots' behavior. For example using these file we can comment on how certain internal states lead to the emergence of different behaviors, or how conflicting motivation influence the robot's behavior. In late experiments in chapters 6 & 7 live print out are used to allow us to engage in appropriate human-robot interaction with the robot.

3.2.2 First experiment: the basic environment

The first experiment was run in the basic environment shown in Figure 3.1, containing two energy (pink balls) and two health (blue balls) resources.

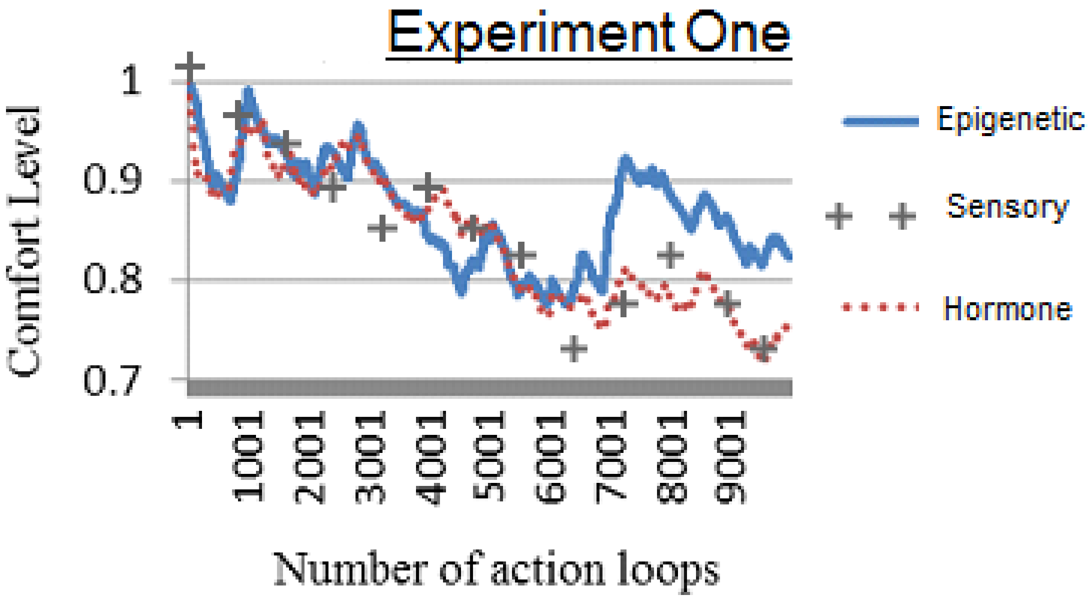


Fig. 3.2 The results of the first experiment

This environment was not expected to provide any great challenges for the different architectures nor to yield any significant variation in behavior or performance. Rather, it was designed as a baseline, both to ensure that each robotic architecture was capable of operating and maintaining homeostasis, and to provide a starting point to assess the adaptive qualities of the epigenetic model in later experiments.

Results of the first experiment

As predicted and shown in Figure 3.2, in this simple environment all three robotic architectures were able to maintain homeostatic balance. However, while the three models were similar in terms of homeostatic maintenance, surprisingly they varied in terms of behavior.

The sensory-driven and neuro-modulated models behaved much as expected: in both cases, when homeostatic deficits were present, they would trigger motivations to move towards and consume the appropriate resources by increase the perceived attractiveness of the resource (see equation 2.8). When not feeding these, the robot would either engage in exploration-like behavior or move to open spaces due to the mechanism discussed in section 2.5 .

In contrast, the epigenetic robot developed and predominantly displayed a guarding-like behavior of the energy resource when not consuming it. This behavior, consisted of the robot either tightly circling or stopping nearby an energy resource.

The cause behind this behavior was a small up-regulation of the E1 receptor and down-regulation of H1 receptor. This change occurred due to a lack of health loss as a result of the simplicity of the environment (i.e., no collisions occurred). The robots' energy of course decrease regardless of the environment, resulting in the minor up-regulation.The T1 receptor also had reduced sensitivity although this did not appear to be an influence in this behavior.

The effect of this combination of epigenetic change, in addition to the lack of other stimuli (due to the relative sterile nature of this environment) meant that this robot maintained attraction and therefore a motivation to move towards the energy resource. However this "basal" motivation was not high enough to overcome the avoidance motivation, leading to guarding behavior occurring when the robot did not need to recover a deficit.

Due to this behavior, the epigenetic robot spent considerably less time exploring the environment in comparison to the other models. As the robots have unlimited access to resources and there is no competition in this environment, this guarding behavior as well as lack of movement could initially be thought of as detrimental in terms of action selection. However, in this particular environment, this behavior —while unattractive— may well be ideal.

As the environment is devoid of any form of stimulation, challenge or source of novelty, the motivation and benefits from moving are limited —getting close to the resources. However, movement also increases motor temperature and the potential for collisions. Therefore, movement here is more likely to lead to an overall increase in homeostatic deficits. By adopting this guarding behavior, the robot essential has "bypassed" the need to manage health and temperature while always being close to an energy resource to replenish its energy levels, thus effectively turning the three-resource action selection problem into a single-resource problem.

In addition, the movement and exploration of the epigenetic model significantly decreased over time in comparison to the other two models, suggesting that as the epigenetic robot familiarizes itself with in the environment, it adapts to the lack of benefit of movement. Moreover, as can be seen in figure 3.2, while all homeostatic levels are close, the epigenetic model did still maintain slightly higher overall values particularly in the latter stages, again supporting the idea that in this case lack of movement is adaptive.

Fig. 3.3 An example of the light movable objects blocking a resource in the second experiment

3.2.3 Second experiment: light movable objects

For the second set of experiments, we added light movable objects to the basic environment. These objects were placed both near the resources and dotted around in the center of the environment, as shown in figure 3.3.

The inclusion of these objects effectively creates a number of new challenges for the robots;

- Firstly, the robot's view of the resources could potentially be blocked, requiring exploration in order to discover them.
- Similarly, the robot may no longer have direct access to some resources, needing to navigate or even push obstacles out of the way to reach them such as in figure 3.3. Since these objects are movable, there is also the option that the robot could knock them over, creating additional challenges.

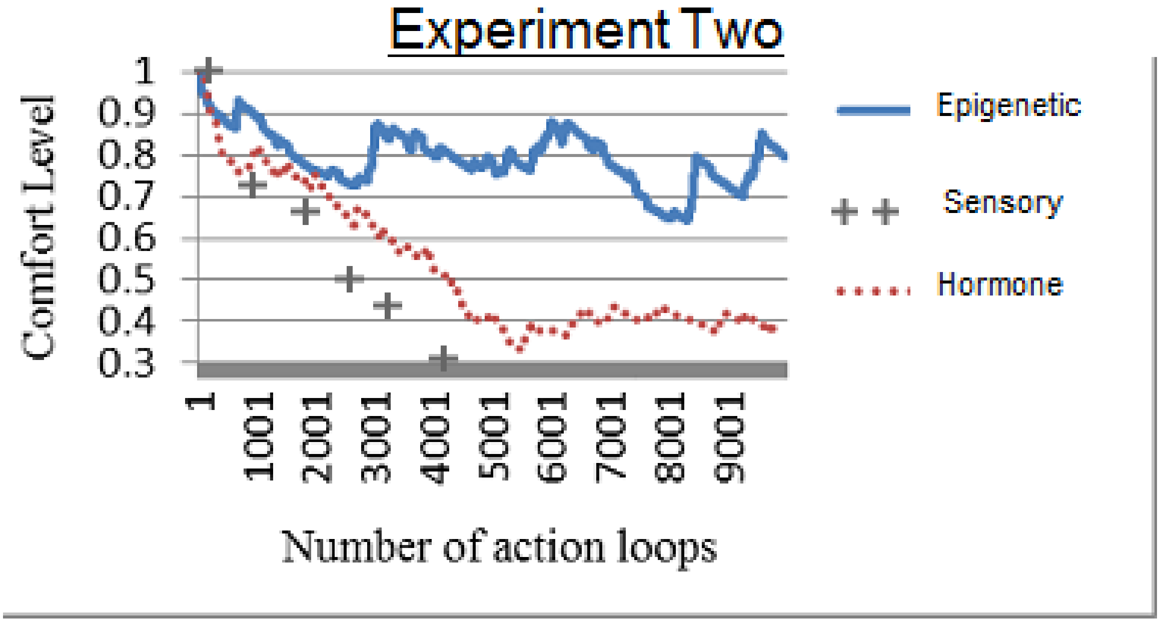


Fig. 3.4 The results of the second experiment

- Finally, the presence of these movable objects greatly increases the likelihood of collisions and therefore increases both the need and the urgency to recover from health deficits. Note that any objects "pushed over/away" from their original location would be restored to their initial location by the experimenter once the robot had moved out of the immediate area.

In the previous experiment (see section 3.2.2), the epigenetic robot developed the previously discussed "guarding" behavior that allowed them to significantly reduce the environmental challenges by effectively focusing on a one-resource problem. However, in this second and the remaining experiments, this type of strategy did not develop, and indeed it would no longer be viable. In this second environment, for instance, in order to reach the energy resource, the robot may need to push past an obstacle, resulting in damage and therefore the need to replenish health. This will therefore require the robot to move between resources which will generate heat. Ensuring that all three variables will consistently be effected by the environmental challenges.

Results of the second experiment

Under this scenario, the non-epigenetic robots performed particularly poorly. As can be seen in figure 3.4, in all cases the sensory-driven robot died in the early stages of all experiments. While the neuro-modulatory model performed slightly better, with only 3 early deaths overall, its performance was still poor.

The inability of these architectures to adapt to the harsher environment stemmed from two main causes. Firstly, these robots often failed to maintain enough health to willfully push their way past the obstacles to reach the resources (i.e., health deficits resulted in these robots trying to avoid collision), and they would instead look for alternative and potentially non-existent paths to the resources.

Secondly, the robots were also unable to adapt to account for the increased time that it took to physically find and reach the resource due to obstacles and their partially hidden nature. This potentially could lead to the robots either dying due to running out of energy, or when deficits reached critical levels attempting a fateful last ditch attempt to push past objects to reach a resource.

However, as the robots did not previously maintain an appropriate level of condition, death was likely to occur during these last ditch attempts. In the cases when the robots survived, condition levels were often so critically low that the slightest collision would result in death.

In comparison, the epigenetic model was able to adapt to the new environment due to a chain reaction of internal changes.

- Firstly, observation of resources which were visible but inaccessible, led to persistent high levels D1 hormones and therefore an up-regulation of the D1 receptor. Leading to increased dominant behavior because of the heightened sensitivity to D1 concentrations.
- The increase in dominate behavior saw the robot more likely to push past obstacles rather than go around. This behavior resulted in increased health deficits, and subsequently increased secretion and concentrations of the H1 hormone. This in term stimulated the up-regulation of the H1 receptor. Resulting in higher sensitivity to the H1 hormone and health deficits. This led to increased motivations to replenish health rapidly emerging after pushing past an object and subsequently the robot maintaining high levels of health overall.

These epigenetic changes ensured that these robots not only would push past object but also maintained enough health to do so without risking critical deficits.

In addition, due to the increased sensitivity to the D1 hormones, these robots developed a more forceful or intense pushing behavior. In this environment due to the way health loss was handled (see equation 2.1), more forceful pushes would push past the light objects faster and tended to result in lower overall health loss (by around 12%) than slower pushes.

The slower pushes often used by the neuro-modulatory and sensory-driven robots also ran a greater risk of objects falling on top of them, leading to further damage; this lead to the direct death of the sensory-driven robot on two cases.

Increased intensity also has the secondary benefit of increasing the likelihood of the robot fully completing the pushing behavior, such as pushing completely past an object rather than giving up and moving away part way through.

The way in which the robots affected the force of their push is simply by driving faster or slower into an object, with the speed of the push dependant on the T1 and D1 hormone concentrations. With D1 increasing speed and T1 suppressing speed (see equation 2.5& 2.8 respectively).

During this experiment the average speed of the epigenetic robots while trying to push was 230mm/s, in comparison to a 150mm/s of the neuro-modulatory and a 60mm/s of the sensory driven robots. Due to the direct dominant behavior of the epigenetic robot resulting in better management of health and energy needs, these robots were also afforded more time to manage any instance of overheating that occurred.

A final note is in regards to “accidental collisions”: The robots here have no learning capacity nor are they able to identify what exactly they are pushing past to reach the hidden resource. Rather the robot can see the resource and detect an obstacle in their-path. As in most cases the obstacle was a tin can, the epigenetic robot adapted to pushing past. However, a possibility of an issue arises if a reflection of the resource (a shiny colored ball) was seen on the border. In these situations (which occurred very seldom), the robot will try to push against a solid object with no give, resulting in a rapid loss of health, in order to try and reach the reflection (which the robot identifies as a resource). Luckily, and highlighting the benefits of a homeostatic regulatory system, this behavior would be abandoned after around 3-4 seconds when the health deficits through the hormone proxy would trigger the avoidance behavior, leading to the robot abandoning the attempt.

3.2.4 Third experiment: moving resources

For the third experiment, we tested the robots in an environment containing 4 moving resources: 2 energy resources and 2 health resources. These resources moved around the arena in a continuous pattern shown in figure 3.5, at a constant speed that was slightly faster than the robot's average speed. At the end of each movement path (represented by a letter in the figure), the corresponding resource would pause for a period of 2 seconds. In order to achieve movement, each resource was attached on top of a e-puck robot which was programmed to follow the paths in figure 3.5. In cases where the (Koala) robot was in the direct path of a e-puck, the e-puck would take the shortest path around the robot before returning to its original trajectory. If there was no valid path for the e-pick —for example if it was blocked in a corner— then it would remain stationary until it was able to resume its movement. At the start of each run, the e-puck were located at different opposite points, e.g., A and E, F and B, H and D, or G and C which we determined randomly before the start of each run.

This environment provided the robots with two distinct new challenges.

- Firstly, the need to develop a consumption behavior suitable for moving resources. Secondly, this environment presents the first situation where the robots can be actively damaged by other elements (objects or organisms) of the environment.
- Secondly, while, as previously stated, resources will attempt to move around the robot if it is directly in their path, they still have a high chance of collision. Therefore, the robots will also need to adapt to co-existence with the resources.

Results of the third experiment

As can be seen in Figure 3.6, the epigenetic robot performed at a higher level overall.

While their where a few occasion of the robots being in the right place and the right time in order to catch a resource with ease (For example finding a desired resource while it was in a corner (i.e., point A or C in figure 3.5) was easier for the robot to catch since it could block its path when approaching, suppressing the resources movement). In most cases the robots would actively need to try and catch the different resources. In order to do so, each model developed distinctive behavior.

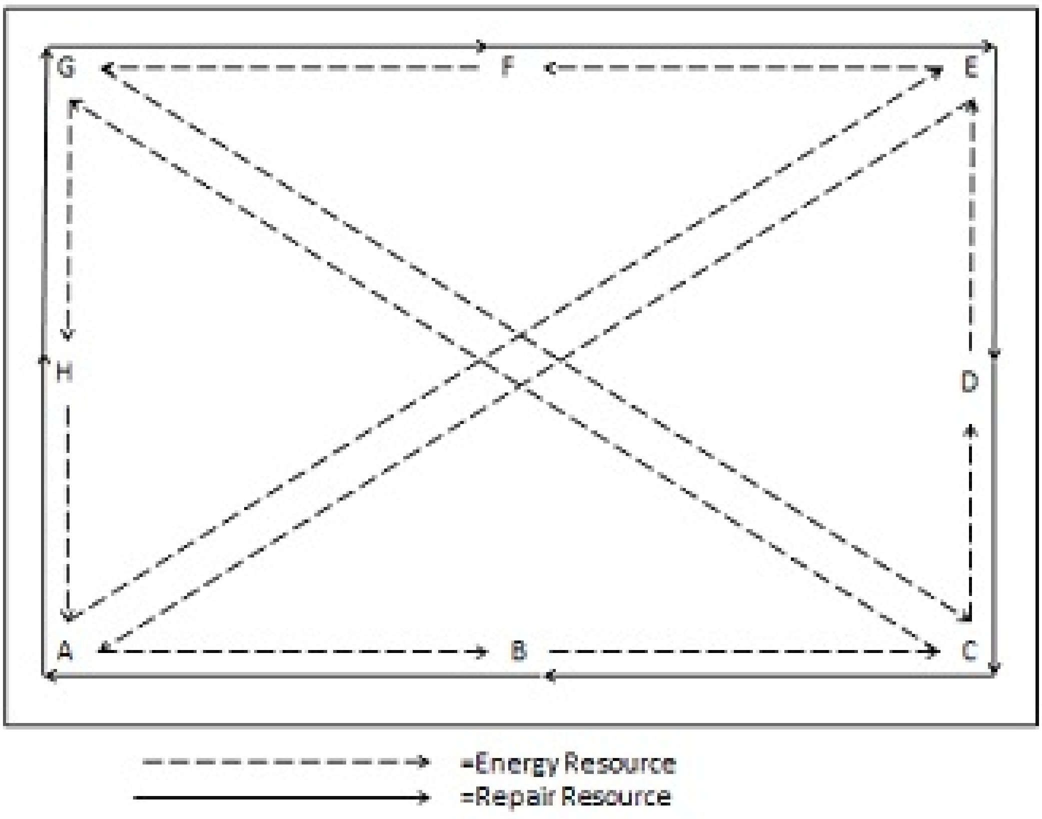


Fig. 3.5 Patterns of movement of the resources (on top of e-pucks) in the third environment. Different letters represent different movement paths. The path of the resources were predetermined and moved by experimenter

The epigenetic model developed an ambush-like strategy, where the robot would remain sedentary until a needed resource passed closely, at which point the robot would pounce, giving chase at full speed. This behavior emerged due the following;

- Firstly, due to initially trying to "chase" the resource, the robot often overheated and secreted high concentrations of the T1 hormone, resulting in up-regulation of the T1 receptor and subsequently increased sensitivity to its temperature
- The heightened sensitivity to it temperature increased the robot's motivation to remain sedentary
- This meant that the stimulation to move to a resource needed to be greater than usual. So while the needed resource could be seen, which importantly lead to increased D1 concentration and subsequently up-regulation of it receptor, at this stage simply being

seen would not provide a great enough stimuli to move towards it. However, if the robot had a need for the resource, the stimuli could provide enough motivation for the robot to “track” it with the robot slowly turning to keep it in view (see formula 2.8)

- However, as the resource moved closer and the stimulation increased the robot would reach a tipping point, where the motivation to move to the resource would outweigh the motivation to stay sedentary. Due to the D1 concentration and up-regulation of the D1 receptor, when the robot moved, it moved at speed.
- Over the next few attempts this behavior was refined as the regulation of the D1 and T1 receptors found ideal levels given the current environment. In the end, resulting in a resource needing to be within around 12cm before the robot “pounced” or gave chase.
- The previously described avoidance behavior (see section 3.5) helped this ambush behavior by ensuring the robot did not obstruct its view on the environment by facing a border at point blank range.

Due to the speed from the epigenetic robot when it gave chase, this chase often resulted in the resource being driven back and pinned against a wall, making it easier for the robot to feed.

This ambush-like behavior had a couple of benefits in this environment. Firstly, by staying stationary, the robot was able to maintain a low temperature, giving it the opportunity to engage in the said bursts of speed without the worry of overheating. Additional as mentioned these bursts of speed provided the robot with a method to catch and pin the resource with out the resource being able to avoid the robot. Overall this behavior meant that, when the epigenetic robot attempted to catch a resource, it had a success rate of 87%, with the average chase taking 4 seconds. almost all failures occurred due to the robot missing the resource during its “pounce” (i.e overshooting the resource) and subsequently losing track of it.

On a few occasions when a resource was not found via the ambush behavior (here this was always due to poor positioning of the robot) the growing E1 and H1 hormones concentrations from any deficits would eventually lead to the exploration like behavior emerging. However exploration and movements without a visible resource would be rapidly suppressed due to rising T1 hormones. This short period of exploration however resulted in the robot reposition itself for another “ambush”.

In comparison, the neuro-modulatory model also had a high success rate of 72% but with a chase taking 14 seconds on average. In comparison to the epigenetic robot, the neuro-modulatory robot simply engages in long chases, often catching the resource when it turned

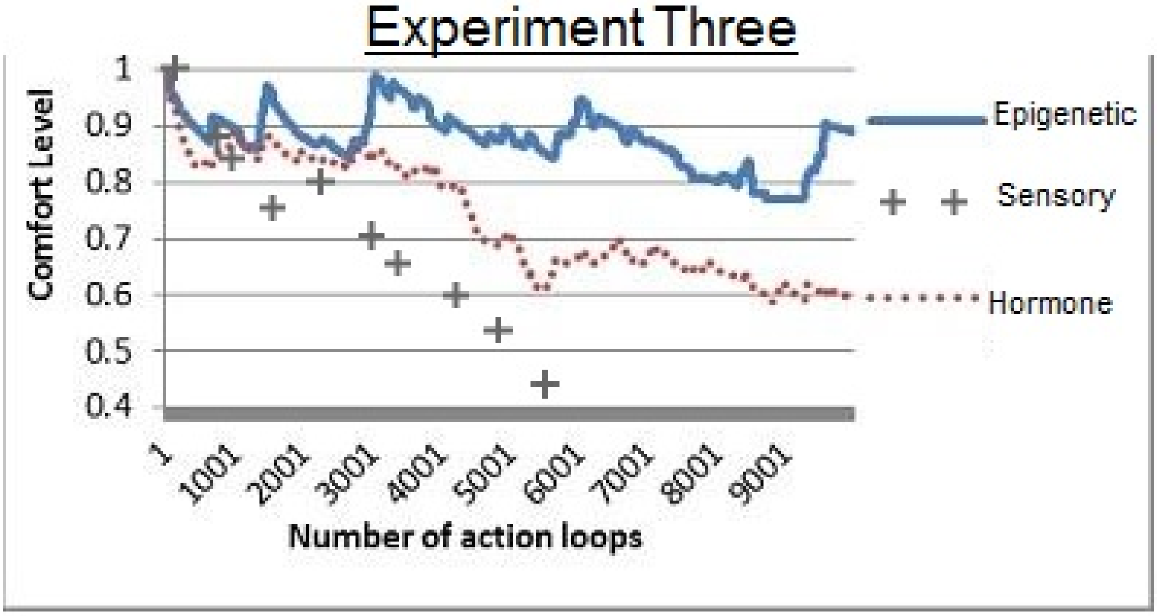


Fig. 3.6 The results of the third experiment.

in a corner. During these chases the robot's speed would increase overtime the closer it got to the resource (greater stimuli) and the longer the chase went on (increased D1 concentration). Failure here was always due to the robot overheating before it reached the resource and needing to suspend movement.

Finally, the sensory-driven model only had a success rate of 13% with an average chase taking 12 seconds. This robot would simply follow a needed resource, however as it could not increase its speed above normal levels (due to the lack of the D1 hormone) and therefore could not catch up with a resource. Therefore unless the robot happened to be near one of the end points where the resource stopped, giving the robots a brief opportunity to consume, it was never successful in catching a resource. In most cases, unsuccessful chases ended due to either the robot getting distracted by another resource or simply losing track. The inability to catch resources meant that these robots constantly had critically low levels of the homeostatic variables and in all cases died before the end of the experiment.

Due to its low success rate, the sensory-driven robot also spent a larger amount of time engaging in chases — 67% of its time in comparison to 7% for the epigenetic and 28% for the neuro-modulatory robots.

Fig. 3.7 Cyclic temperature changes in the fourth experiment.

3.2.5 Fourth experiment: a dynamic climate cycle

In the fourth experiment the robots were placed in an environment with a dynamic climate, where the standard ambient temperature would increase and decrease cyclically over time, simulating a day-and-night temperature cycle. A full cycle consisted of twelve 10-second periods for a total two minutes, as can be seen in Figure 3.7.

To simplify the model, ambient temperature was ranked on a 0 (cold) to 10 (scorching heat) scale. At peak temperatures, the robot would only be able to sustain its average speed for a period of around 5 seconds before it would overheat. At this point, even slow movement would be unsuitable. In order to increase the dynamics of the environment, temperature was allowed to fluctuate by up to 2 points to simulate potential meteorological phenomena. These fluctuations were calculated at the start of each 10-second period and lasted until the next period.

While these meteorological phenomena do result in fluctuations in climate, generally speaking the peak environmental temperature is reached after every 2000 steps.

In this environment, the robots were presented with two main challenges.

- Firstly, they ideally needed to address any homeostatic deficits during the cooler periods so that they may lay dormant during the high heat.
- Secondly, due to the dynamics of the climate, they needed to be able to quickly adjust to changes in the environmental conditions. The results from this experiment are shown in Figure 3.8.

Results of the fourth experiment

As can be seen in this figure, all the sensory-driven robots died before the end of the second climate cycle, with causes of death evenly split between lack of energy and overheating. Simply put, these robots could never adapt to the need to take advantage of the colder period; therefore, during the hot period they were left with the option of either moving to a resource and overheating or staying still and running out of energy.

The neuro-modulatory model fared better. While it was still unable to fully take advantage of the cold periods, a slow-paced moving behavior emerged during the hot periods. This was

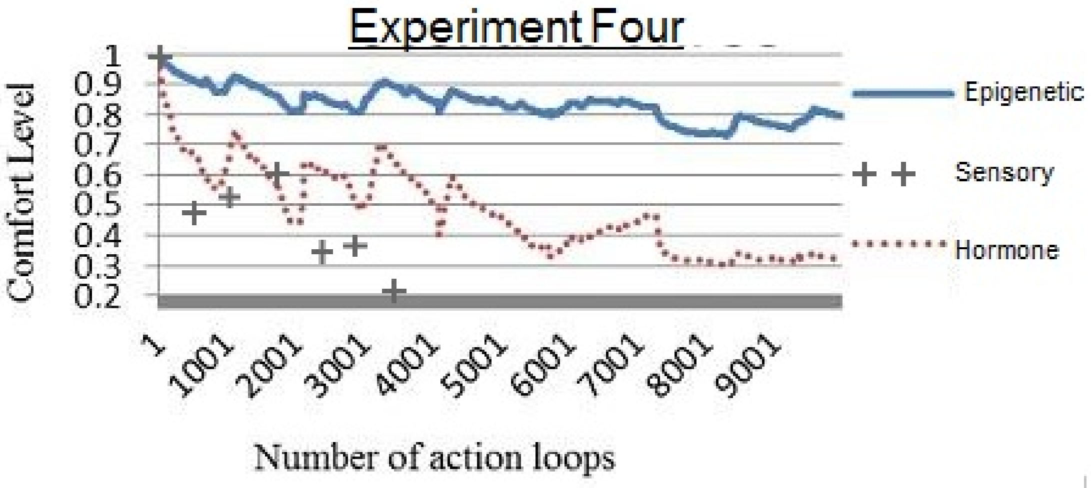


Fig. 3.8 The results of the fourth experiment

due to suppression of speed as a result high concentrations of T1, allowing the robot to move without overheating. Thus giving it a chance to reach a resource during this period if it was needed. However, even with this behavior, three deaths were still recorded due to energy loss. Out of these three deaths, two of them occurred during one of the “meteorological phenomena” when a temperature spike occurred.

Finally, the epigenetic robots developed two contrasting behaviors in order to survive the periods of high ambient temperature. One group developed a “hibernation” behavior, during which the robots became highly attracted to the different resources during the cold period and thus could fully replenish any deficits before laying dormant during the hot periods.

This “hibernation” behavior emerged due up-regulation of the three *eH* receptors ,particularly the T1 receptor. With the T1 receptor becoming up-regulated due the heat during the day. Resulting in the robot completely suppressing movement when it got hot. As the robot could not move during the day, in the early stages any deficit in health or energy could not be sated during this period, resulting in rising concentrations and therefore up-regulation of their respective receptor. This meant that when the period of heat past, and the T1 concentration degraded these robot would immediately seek to sate any deficit in health or condition regardless of how small. After all needs are sated any time left in the cold period would be spent exploring.

The second behavior that some robots developed was instead to simply stay near the energy source at all times, except when the occasional need to repair arouse; this behavior permitted the robots to continue consuming energy during the hot temperature climate period, with only very limited movement needed. This behavior has a similar basis to the “Hibernation” except only being an up-regulation of T1 and E1

Which behaviors is adopted seems to be associated with the amount of health deficits the robots have during the first few hot cycles. If the robot goes into these cycles in perfect health the second behavioral phenotype is adopted. If however the robot goes into the cycles with some health deficits then the robot adopts the “hibernation” behavior.

Overall both behaviors were effective as neither group suffered any deaths and neither contributed to a significantly higher internal temperature. Their main difference was the amount of exploration of the environment associated with each, with the “hibernation” strategy giving rise to more active robots in terms of exploration. Interestingly in later experiments in more dynamic environments such as those in chapter 6, under similar circumstances robots will only adopt the hibernation like behavior. The reason for this is likely that as the hibernation behavior allows for additional activity and exploration it becomes more attractive or "rewarding" as the robots are exposed to more novelty.

In Figure 3.9 we provide the “maps” of the movement traces or of the three different robots during the first set of these experiment. These maps have been generated using the speed of the wheels of the robots that was recorded during the experiments[1] rather than tracking the robots’ physical movements with an external camera, which was not available to us during the experiments.

As can been seen in Figure 3.9, the epigenetic robot explored a large portion of the environment while maintaining a low temperature. In contrast, the two other robots moved primarily between the two resources located in the top left and bottom right corners. Even though these robots engaged in limited exploration, they had a significantly increased occurrence of high internal temperatures. For the neuro-modulatory robot, these high temperature periods were normally a result of the robot running low on energy and needing to move quickly to replenish. For the sensory-driven robot, however, these periods simply represented a lack of adaptation to the increased ambient temperature, with the robot continuing its normal behavior even in the hot conditions.

---

[1] This method of tracking the movement of the robots might have produced minor discrepancies between the robots’ real movement and those generated in the maps, although it provided a good enough approximation for the purpose of showing the differences among the three robot architectures.

Fig. 3.9 Movement maps of the three different robots. The color of the lines represent the current internal temperature of the robots: light grey represents a temperature between 0-50 units, dark grey 51-75 units, and black 76-100 units.

3.2.6 Fifth experiment: uneven resources

In the fifth experiment we created an environment which challenges the agents by having uneven resource density. In order to do this, three sets of five runs were conducted. The first set featured four additional energy sources, giving a total of five. The second set featured four additional repair sources. The third set featured a hotter climate, which if using the scale from experiment 4 (cf. Section 3.2.5) would be heat of an intensity of around 6 (see figure 3.7). This meant that while the robot could move at its regular speed, it would need to take regular breaks to cool down.

The first two sets of experiments provided similar challenges for the architectures of all the robots. Firstly, the increased stimulation caused by the additional resources could potentially pose a distraction while searching for the rarer resources. Secondly, resources were semi-dynamic and could be moved when touched, potentially blocking the view of the rare resources. In order to be successful in this environment, the robot would need to be more sensitive to resources and act in a timely manner when the opportunity to replenish was present.

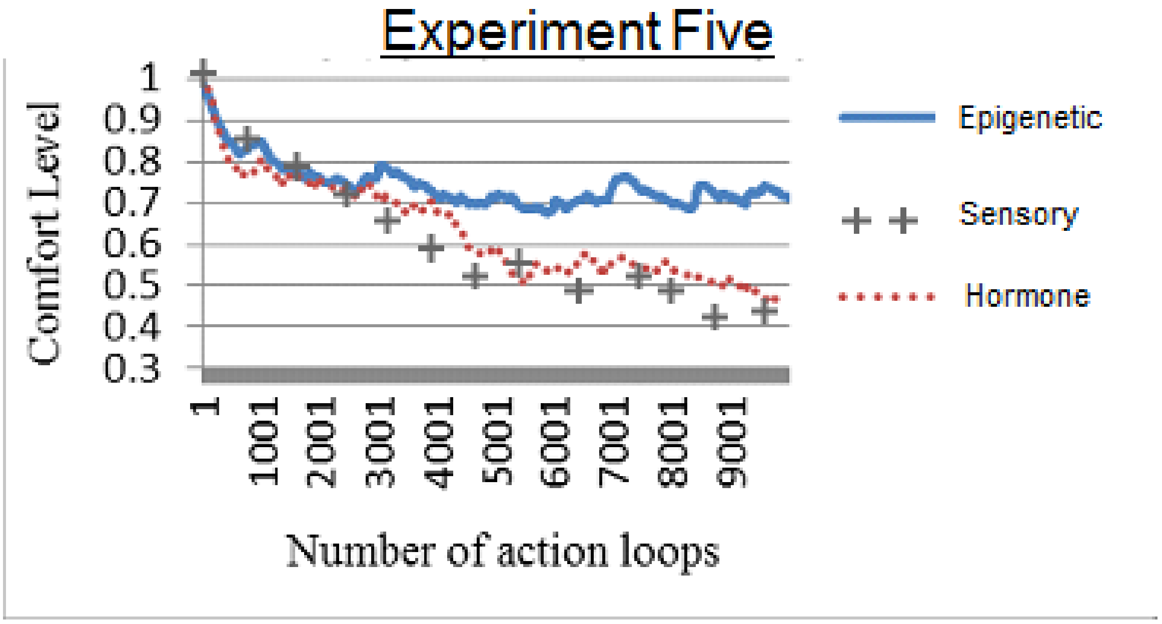


Fig. 3.10 The results of the fifth experiment.

Results of the fifth experiment

Results are shown in Figure 3.10. The sensory-driven model was in all cases able to operate at a low but passable level. While there were four deaths spread across the experiments and cases of low homeostatic levels, this was the only time, excluding the first experiment (Section 3.2.2), when a majority of the robots featuring this model survived until the end. Behaviorally, there is little to report, as the robots here behaved almost identically as in experiment one. In fact, the only real difference was that, due to the unbalanced density of resources, these robots spent more time looking for the rare resources and less for more common ones.

Although the neuro-modulatory model performed better than the sensory-driven model, it still had a hard time dealing with uneven resources. The overload of stimulation from the excess resources meant that the homeostatic deficit correlating to the rare resources was on average 32% greater. The two occasions when the rare resource was hidden behind another resource were also extremely problematic as the robot would consistently stop searching for it in order to satisfy the other motivations.

In contrast, the epigenetic model was able to successfully adapt. Early homeostatic deficits led to the rapid secretion of hormones triggering early epigenetic changes and increasing sensitivity of the associated hormone receptors. This in turn made the robot more sensitive to these deficits, filtering out the distraction and noise of the other resources. Under the epigenetic model, when both stimuli were present, the deficit of the overabundant resource would need to be around 42% greater in order to attract the robot's attention. Further, in the four cases when the rare resource was hidden, the deficit for the common resource would need to drop to between 60-80% before the search for the rare resource was interrupted.

Finally, the set of experiments with the high level of temperature did not yield considerable differences in terms of homeostatic levels between the three models. While the epigenetic model did perform better in the experiments, the biggest difference was qualitative in terms of the behaviors that emerged from the models. The epigenetic model developed a ""stalking-like" approach due to up-regulation of the T1 & E1 receptors. This behavior emerged in similar manner to the "ambush" behavior seen in section 3.2.6 except with a more even level of regulation of the T1 & E1 receptors. This variation was due to the

- The high temperature resulted in increased T1 secretion and subsequent up-regulation of the T1 receptor, resulting in the suppression of movement
- However, unlike the ambush behavior in section 3.2.6; resource did not move, instead the robot needed to move to them. This meant that by staying stationary the robots energy would simply decrease, leading to high level of E1 concentrations and subsequent up-regulation of the E1 receptor would occur.
- The high concentration and up-regulation of both the E1 and T1 hormones led to the robot being modulated by opposing concentrations, which resulted in a slow exploring behavior. However during exploration when the robots sees a resource, when the robot gets close enough, like the ""ambush" the motivation to recover outweighs the desire to maintain a low temperature, and the robot busts forward

This regulation resulted in the robot moving toward the resource taking breaks until it was within a certain range, and at which point it would burst forward at full speed.

In contrast, the neuro-modulatory model would maintain a slow constant speed due to equal levels of the E1 and T1 and move towards the desired resource. While both behaviors worked, the epigenetic approach generated less overall heat, and was therefore slight preferable for this environment.

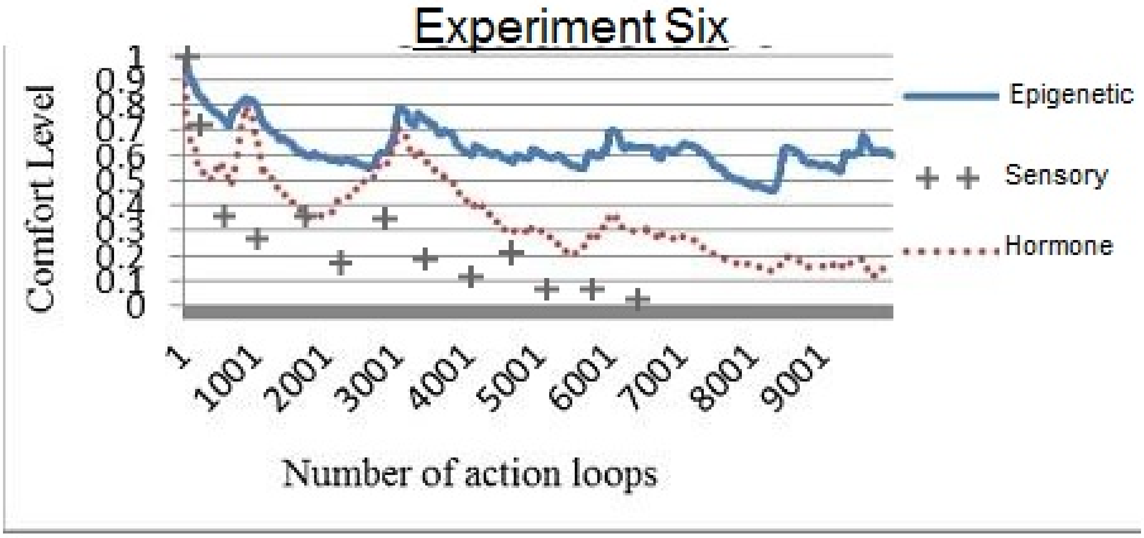


Fig. 3.11 The results of the sixth experiment.

3.2.7 Sixth experiment: temporal dynamics

In the final experiment, we placed the robots in an environment where the energy resources showed the following temporal dynamics: they appeared at specific locations in the environment once every minute, and they remained available for different durations. For the first five sets of runs, the energy source would remain available for 30 seconds before being removed. In the second five sets of runs the duration was reduced to 20 seconds and the final five sets of runs saw the resource only accessible for 10 seconds of every minute. Temporal dynamics was limited to the energy resource only in order be able to examine the robot's ability to deal with the increasing disparities between the availability of the repair and energy sources. It is worth noting that the robots had no capacity to monitor time; therefore, they could not try to directly predict when the resource will appear. Instead of that, the robot will adapt over time to the scarcity and rarity of the resource. A strict time period was used to ensure that each robot had same constraints and opportunities.

Results of the sixth experiment

As can be seen in figure 3.11, the sensory-driven model was unable to deal with this temporal dynamics, never managing to survive a full run. In the first five sets of runs, when the resource was available for 30 seconds, the robot was able to survive around eight minutes on average using the same behavior as in experiment one (see section 3.2.2). However, as the period for which the temporal resource was available decreased, the robot struggled increasingly and by the final five sets of runs all agents had died before the 2-minute mark.

The primary problems for this robot were its inability to adapt to “rare” conditions, as in experiment five (see section 3.2.6), and its inability to adapt its speed in order to move promptly towards the temporal resource. Overall, when a temporal resource was available, the sensory-driven model managed to reach it only 17% of the time on average. This is in comparison to an average success rate of 84% for the epigenetic and 62 % for the neuro-modulatory models.

Comparing the epigenetic and neuro-modulatory models proved a bit more interesting. Both robots performed at a similar level during the first five runs with the 30-second window of opportunity. While the epigenetic robot moved more promptly to resources when they appeared due some up-regulation of the E1 receptor, neither robot was ever in any real danger of missing an opportunity. For the epigenetic robot, up-regulation of the E1 receptor occurred as a result of increased energy deficits and subsequent E1 hormone concentrations, due to the resource not always being available when needed.

However, as the window of opportunity shrunk, the differences between the two models became apparent. Since the point where the resource would appear next was unknown to the robot, it was inevitable that both architectures would now miss some opportunities to replenish. However, the epigenetic model was generally quicker to move to the energy resource due to further up-regulation of the E1 Receptor.

Finally, due to missed opportunities to fully recover deficits, both robots often showed a significant level of the D1 hormone. This in turn resulted in increased aggressive-like behavior and the occurrences of collision in later runs, subsequently increasing the need for the repair resources. In multiple cases this lead to similar or greater deficits in health in comparison to energy. Leading to similar or greater hormone concentrations of H1 in comparison to E1. As Health deficits could and were readily recovered, in the epigenetic robot they did not lead to any significant up-regulation of the H1 receptor.

However, the health deficits did cause the neuro-modulatory robot to sometimes go to the readily available repair resource during the limited periods when the energy source was present and seen. This was due to both hormone concentrations having equal effect on the robot.

In contrast, the epigenetic model, due to the up-regulation and heighten sensitivity of the E1 receptor had effectively adapted to the rarity of the resource. Even with health deficits present, it only missed the opportunity to replenish energy once, when its health levels were critical.

In total, seven of the neuro-modulatory robot runs ended prematurely compared to a single death in the epigenetic model.

3.3 Conclusion

In these early experiments we have found that an epigenetic like mechanism significantly and consistently improves the robot's adaptability in both quantitative and qualitative comparison to the sensory-driven and neuro-modulatory models that we used as comparison points. We found that epigenetic mechanism was indeed able to regulate the internal environment of a robot in relation to its environmental conditions. This mechanism lead to a robot being able to regulate its development in an appropriate manner. With the robot becoming highly adaptive.

In all our experiments, using the previously discussed comfort performance measure (see section 3.2.1), the epigenetic robot maintained significantly higher levels of comfort than the other models. In addition, the variance between the runs in experimental sets of the epigenetic robot was considerably lower, indicating consistent performance.

From the perspective of a more qualitative analysis of the robot behavior, we have also seen how on each occasion the epigenetic robot's performance and behavior were better suited to each of the environments. These qualitative observations included the development and emergence of successful "unplanned" behaviors and tactics such as pushing or hibernation. In addition, as we discuss in each experiment overview we also consider other factors such as exploration and tolerances to deficits. For exploration, in all cases but one —the first experiment, in which the epigenetic robot had reduced levels of exploration due to the barren nature of this environment and the lack of challenges— the epigenetic robots explored significantly larger portions of their environments.

In the case of tolerance, in all cases the epigenetic robot developed tolerance and therefore developed goals and motivations that were more in tune with its environment in comparison to the other models —for example in a environment with plentiful energy supply the epigenetic robot would tolerate larger energy deficits.

Additionally since we have essentially tested different stages of the robotic model we can see that the epigenetic mechanism, the only difference between the neuro-modulatory and epigenetic architectures must play a critical role in what we have observed. This not to suggest that the other aspect are unimportant, far from this as we will discuss in later chapters it is the combination of all three components that bring about this adaptive capability that observed .

Due to the significant adaptive advantages that a simple epigenetic like mechanism seemed to have provided an autonomous robot with we are confident that it warrants greater investigation and in the following chapters will begin to investigate this mechanism in both more complex architectures and environments settings.

Chapter 4

An artificial HPA-HPG axis for curiosity-driven motivation

As we have shown in our early experiments a mechanism inspired by epigenetic principles allows a sensory driven autonomous robot to rapidly and successful adapt to a range of different environmental conditions. One question that came up during these experiments, and the focus of this and subsequent two chapters was if the hormone like chemicals could regulate simple development in an appropriate manner given the robot's environment, could they also be used to regulate other aspects of developmental plasticity i.e., learning, cognition and neural growth?

Learning has so far not be examined in our experiments. The previous experiments in chapter 3 all took place within what could be considered relatively controlled, sterile and stable environments particularly when compared to real world conditions.

While their were certain dynamic elements which only the robot possessing the epigenetic mechanism was able to adapt to, the fundamentals of the environments remained constant, e.g the blue ball would always replenish energy, or getting to close to an object would result in a crash and a subsequent loss of health. Due to both the constant aspects and general simplicity of these past environments, our robot could be provided with highly accurate environmental information with limited effort from the developer.

However, in a truly dynamic real world environment designing a robot with complete environmental knowledge would likely be impossible given both current hardware limitation, and the fact that we as designers are unlikely to possess this level of knowledge. To further complicate the issue we must also consider that the pre-programed environmental information must both be relevant for the robot's characteristics and flexible so as to still be usable if these characteristics change, e.g a sensor or component of the robot became damaged.

If a system is designed which relies on imperfect or incomplete environmental knowledge, even slight changes in the environment can lead to significant and often unpredictable changes in the trajectory of the robot's behavior [125, 15, 132, 86]. While environmental changes tend to modify the organism's behavior in relation to the environmental change (see e.g., [29, 32, 149]), significant changes to the environment of robots possessing pre-programed/determined adaptation mechanisms can lead to behaviors that are not only unsuitable but may render the robot inoperable or even lead to potentially dangerous behaviors given the vast number of potential environmental variation in the real world [140, 70].

Before we begin to look into the potential for our system to be able to regulate learning in an appropriate manner the robot must first have a drive to interact and discover. This is to say we suggest that the robot needs an intrinsic mechanism to be able to identify what aspects of the environment to learn about. Here we expand upon are past mode to introduce the idea of Novelty, generated using additional artificial neurohormones. The purpose being to regulate novelty seeking behavior to be appropriate based upon current conditions.

In this chapter we will investigate if neurohormones can be introduced to create desired novelty seeking behavior, which we suggest is needed in order to introduce neural development into the model.

4.1 Novelty-driven learning

In order to operate in a real world dynamic environments it would seem as if our epigenetic mechanism needs to be expanded in order to include some form of explicit learning capacity, allowing it to adapt to new and changing environments, both autonomously and in relation to its own morphological and physical capabilities in a similar manner to animals [138, 117, 112]. However, we must consider how to implement such a mechanism.

In animals learning is not a passive phenomena which simply occurs. In particular in humans infants learning is driven by their own activities, regulated by intrinsic motivations to create learning opportunities and experiences [113, 53, 78].

A well documented example of this process is discussed by [113], where he demonstrated a learning scenario by placing a rattle in an infants hand. Initially the infant would simple hold the rattle until by chance they found that by moving the hand holding it, they could cause the rattle to make a noise arousing the infants interest. This arousal is then said to be responsible for the infant to deliberately recreate the arm movement intentional of making noise [137].

At some point, after the infant has discovered a range of variation and patterns, and the outcome of shaking the rattle becomes predictable, the infant loses interest in the rattle. This loss of interest after throughly exploring the rattle highlights the potential underlying role that curiosity and novelty may play in the learning process [78]. To this extent it has been suggested that the infants' exploration of rattles and similar objects are instances of curiosity learning and potentially reflect a intrinsic motivation that select actions which lead to perceived “interesting and novel” sensorimotor outcomes [113, 78, 53].

4.2 Novelty-driven learning in AI

As we discussed in the beginning of this chapter, in order for our robot to adapt to more dynamic environments a learning system is likely needed to allow the robot to discover and overcome new and unexpected condition. While a range of different of learning mechanism have been demonstrated, these systems tend to be desgined and tuned around certain explicit goals and environments (see [135, 70] for an overview).

This tuning can be either very direct such as predetermining the weighting of environmental cues, or more subtle through the use of mechanisms such as reward feedback, fitness functions and activity functions [70]. The problems with these types of mechanism though is that they many fail to offer a truly open ended and ongoing learning experience for a robot.

However, an approach based upon novelty driven sensory motor learning seen in animals may well provide the open ended system required for adaption to real world environments. Under this type of approach, like with animals, novelty and curiosity could potential be used to encourage appropriate desired interaction with the environments, a sensory motor based learning system will then permit the robot to lean about different aspects of it environments as a result of the novelty driven interactions.

This system therefore consist of two main aspects, novelty driven motivations and a sensory motor learning system. Here in this section will discus and outline our approach to a new novelty mechanism which achieved by expanding our hormone driven epigenetic system. Later in chapter 5, will introduce the learning system which we called an Emergent neural network or ENN.

Before we go into greater detail of our novelty it is prudent to briefly overview novelty driven mechanism. In the AI community the idea of using novelty and curiosity as a way to facilitated learning in autonomous roots has been explored in numerous environments and situations [110, 99, 2, 109]. It has been suggested by [110] that the research into these novelty driven learning mechanism can be separated into three main approaches

- Error Maximization: In these model of novelty based learning and exploration, the robot will choose actions which have the largest error predictions see [138]. While these models prove operable in sterile environments, as [110] suggests, in more dynamic real word situations where noise is present and both non deterministic and non- homogeneous these models often become trapped in a behavioral loops which may be unsuitable for the current environment.
- Progress Maximization: These model attempt to address the issues of the Error maximization approach through the introduction of a mechanism called the "knowledge gain assessor". This mechanism works by evaluating the difference between the expected mean error in the recent future in comparison to the expected mean error of the recent past. This evulation is designed to safeguard against the robot becomes stuck in behavioral loops. However, this approach has only seen limited implementation such as [126] and primarily has only seen use in single action robots [110].
- Similarity-Based Progress Maximization: Finally a similarity-based progress mechanism seeks integrate curiosity based learning into a developmental robotic platform. In this type of approach rather than compare the mean error between recent past and future, the robot will instead compare the error between similar situations [110]. This type of mechanism seem so far to be the most successful approach to curiosity driven learning

However, while we use a novelty based approach to facilitated learning in our model, our approach is different from the above. Firstly in our model the attraction to novelty arises due to the interaction of hormones (see section 5.2.3 and the nodes in the ENN. Specifically the different hormone concentrations modulates the network in away that leads to neural pathways associated with novel stimuli,environment or objects becoming either more active or suppressed depending upon the robot's internal state and past experiences. Secondly as discussed in section 4.3 while the novelty value of an object is partly determined by the "error" or uncertainty associated with an object, other factors also influence the robot's perception.

In comparison to other curiosity based systems two additional points need to be made.

- Firstly, we maintain the previous survival related homeostatic variables, which means stimulation and motivation to investigate and interact with sources of novelty will compete against the robot's other motivations such as avoiding dangerous environments or recovering a homeostatic need. One of the benefits of multiple motivations is that it stops the robot becoming stuck in a behavioral loop as it perception of said novelty will be constantly changing both as it learns about the object and as its internal state changes.
- Secondly, the robot's attraction to novelty under this mechanism is dependent upon combination of the size of the novelty, past experiences and its internal state. For example exuding any past learning which may effect the robot's perception of novelty, we can expect a robot which perceives itself as being close to "death" will likely avoid any source novelty or uncertainty preferring instead the safety of the known, in contrast a robot which perceives itself as being in a "moderate" state may be attracted to smaller novelty or uncertainty but still will avoid larger novelty sources. Finally a robot which perceives itself as being in a good state will likely be attracted to larger novelty sources.

4.3 Hormones and novelty

In order to create a novelty based motivation mechanism that would permit our robot to interact appropriately with source novelty in its environment, we expand upon our hormone system by adding two new artificial hormones called stress and curiosity, functionally akin to two chemical modulators, the steroid hormones corticosteroids and testosterone respective. These two hormones, stress and curiosity, replace the previous dominance (D1) hormone leading to a total of 5 hormones in the model, the three *eH* hormones E1, C1 and T1 and the two new *nH* hormones.

The idea of utilizing artificial variations of these two hormone to regulate novelty seeking behavior lies in the role of these hormones in biological organism. In biological organisms both steroid hormones have long drawn particular interest for their role in modulating a wide range of value-laden survival and social behaviors. This occurs due to the interaction between the hormones and one of their primary targets, the amygdala [69].

While the exact mechanisms are unknown, once the hormones have reached the amygdala, their behavior is better understood. Testosterone is linked to promoting outgoing reward-seeking behaviors such as dominance, aggression, exploration, and curiosity [34, 89]. In contrast, corticosteroids are related to avoidance and withdrawal behaviors [19, 102]. Moreover, these hormones do not only modulate emotional processing of the amygdala but are also believed to affect its neural connectivity to other areas of the brain, particularly the orbitofrontal cortex. Exposure to corticosteroids leads to strengthening these neural connections, while testosterone weakens them [97, 143]. As the orbitofrontal cortex is associated with decision making, strengthening or weakening the emotional input from the amygdala could result in additional behavioral modulation [12]. Although the two steroid hormones have significant potential to modulate behavior, actual studies into the individual roles of these hormones do not always offer conclusive evidence. This is particularly noticeable in human studies where results are normally limited to observation of subjects, which can even be contradictory. There are at least two likely explanations for this.

- Firstly, there is significant evidence to suggest that both cortisol (CHT) and testosterone (T) work in tandem to modulate behavior and it is the ratio or imbalance between both chemicals that is important (Montoya, 2012). For example, in a situation with a high T/CHT ration (high T low CHT) aggression is more prevalent than in a situation with an equal ratio, even when the T level remain constant (e.g.,[115]). Therefore studies which simple focus on the level of a single hormone may not be showing the entire picture.

- Secondly, the effects of hormones and secretions are likely to be subjective to the individual . This is particularly relevant for organism with highly complex cognition and neural mechanisms, such as humans. Where aspects such as learning, planning, normative behaviour and beliefs gained through life-long experience will affect “consciousness” and therefore can lead to differences in individual emotional processing [4, 75].

However, it is not only the level of neural mechanisms that can lead to subjective hormonal modulation. More recently, evidence has arisen of prototypical plasticity in the neuroendocrine systems responsible for the secretion and regulation of T and CHT. Changes in gene expression occurring within these neuro-endocrine systems are known to be associated with extreme forms of behavior [95] .

However, it is also highly likely that these changes could have an effect on day-to-day behavior. The neuroendocrine systems of T and CHT consist of the Hypothalamic–pituitary–gonadal axis (HPG-Axis) and the Hypothalamic–pituitary–adrenal axis (HPA-Axis) for T and CHT secretion respectfully. While these two axes are often considered separate entries, they are interconnected through feedback loops. Specifically, research has shown that the HPA-Axis suppresses the activity of the HPG-axis on all levels (see figure 4.1). In addition the HPA-

AXIS contains a negative feedback loop that consists of glucocorticoid receptors which, in response to rising corticoid levels, signal for the suppression of the axis activity [102].

However, this is not a simple static relationship between cortisol levels and HPA activity. Research has suggested that the glucocorticoid receptors responsible for the feedback are susceptible to epigenetic changes consisting of up and down regulation. Down-regulation, which is a reduction in the total number of receptors, leads to reduced sensitivity to corticoids and thus weakens the negative feedback loop. In contrast, up-regulation leads to an increased number of receptors and therefore increased sensitivity and a more reactive negative feedback loop [77, 95, 149]. Down-regulation of glucocorticoid receptors has been linked with, and believed to be triggered at least partially by, continuous high levels of corticoids in the system [95]. Up-regulation, on the other hand, has been associated with positive upbringing and experiences during early life with dopamine considered a potential chemical trigger [77].

So as can be seen above, both corticosteroids and testosterone not only potentially play a critical role in modulating value-laden behavior such as curiosity and reward seeking but also potentially have an important role in regulating each other.

4.4 Implementing an artificial HPA/HPG axis

In order to investigate if corticosteroids and testosterone can be used as a mechanism to generate and regulate curiosity we designed an implemented a simplistic artificial HPA/HPG axis which was tested in section 4.5. It should be stressed that this system does not try to replicate the complexity of the HPA-HPG axis which in consist of a multitude of different components and chemical which vary between both species and gender. Rather we design our artificial axis based on the following principles

- The stress hormone (corticosteroids) is secreted in relation to perceived stressful stimuli. The Curiosity hormone (testosterone) is secreted in relation to perceived positive stimuli.
- The Stress hormone should reduce novelty seeking behavior while the curiosity hormone should increase it.
- The stress hormone should suppress the secretion of the curiosity hormone (we test the importance of this later in section 4.5)

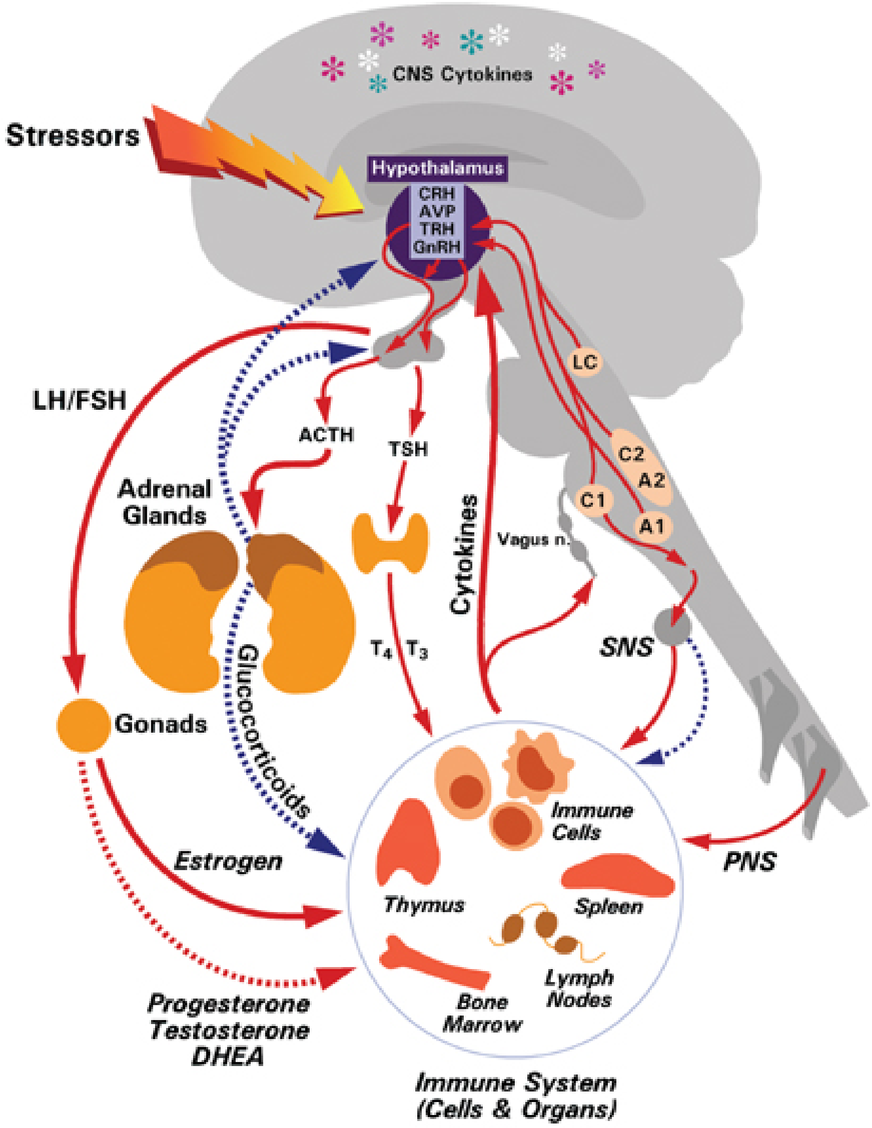


Fig. 4.1 The potential interaction of the HPA-HPG axis, where Corticosteroids are believed to suppress HPA and HPG axis at all stages. It should be noted that differences in the interaction between the axis may be present due to gender differences. Dotted lines represent negative feedback loops, and solid lines represent positive feedback loops taken from [87]

- The sensitivity of receptors to the hormones is dependent upon past internal exposure (we test the importance of this later in section 4.5)

Given the above in our model the secretion of the neuro-Hormones stress ($nHSecretion_s$) and curiosity ($nHSecretion_c$) is given by equations 4.1 and 4.2 respectively.

$$nHSecretion_s = roD \times oS \times nS \tag{4.1}$$

where roD or the "perceived risk of death" is the sum of all homeostatic deficits, oS or "overall stimulation" is the sum of the total amount of stimulation (regardless of its type), and nS is the sum of perceived "negative" stimuli. By "negative stimuli" we refer to the stimulation associated (by the robot's neural network) with the worsening (i.e., the increase of the deficit) of a homeostatic variable. In other words, negative stimulation nS is the sum of any output associated with the worsening of a homeostatic variable, and is calculated by the synaptic function of the output nodes of the neural network, as shown in equation 5.5. The overall stimulation oS is also determined by the synaptic function of the output nodes of the network and is the sum of the total synaptic output.

$$nHSecretion_c = \frac{pS + \sum_v r_v}{nH_s} \tag{4.2}$$

where pS is the sum of all perceived "positive" stimuli, $r_v \geq 0$ is the (perceived) recovery of a homeostatic variable v during the current action loop, and nH_s is the concentration of the stress hormone which suppresses the secretion of nH_c. By "positive stimuli" we refer to the stimulation associated (by the robot's neural network) with the recovery (i.e., the correction of the deficit) of a homeostatic variable. In other words, positive stimulation pS is the sum of any output associated with the recovery of a homeostatic variable, and is calculated by the synaptic function of the output nodes of the neural network, as shown in equation 5.5.

Once secreted, both of these hormones decay at the same constant rate as the other hormones as shown in equation 2.4. In addition like with the other hormones, the sensitivity of receptors associated with each nH is modulated using the same epigenetic mechanism as used previous for the other hormone receptors see equation 2.10.

These new hormone addition to the model along with the removal of nH_{D1} see model previously shown in figure 2.4 changed to the new architecture shown later in figure 5.4

Table 4.1 The types of objects the robot can detect using our simple classification network along with the predetermined fixed novelty and novelty scaling values

Number of IR sensors active	Type of object	Initial Novelty	Stress Scaling	Curiosity scaling
0	Empty space	1	-1	1
1	Small object	200	-4	2
2	Medium sized object	-200	-8	8
3	Large object	-400	-12	14
4+	Wall	0	10	0
0, but neighboring sensors active:	A hole or gap	-800	-16	20

4.5 Experiments

In order to test if the proposed addition to the hormone architecture could lead to the generation of appropriate curiosity within the robot a simple 2 part experiment paradigm was designed. However, we have not yet introduced the ENN which these new hormones regulate in order to generate the desired curiosity driven motivation.

The reason for this is here we wish to see if the hormones can lead to the generation of the desired motivation rather the actual learning outcome. In order this achieve this for this experiment we use a simple classification network in addition to the previous sensory motor architecture rather than the ENN. By doing so we are to focus exclusively on the robot's ability to generate appropriate of curiosity driven motivations.

This new classification network allowed the robot to detect 6 types of object depending upon IR sensor activity as shown in table 4.1. Each type object the robot could detect was give an initial fixed novelty value as well as scaling depending upon the hormone concentration, again show in table 4.1.

The initial novelty value as well scaling values where based on a simple principle, the bigger the object was, the more novel and interesting but also potentially more stressful or "scary" it became. Therefore a large object for example would require a high concentration of nH_c and a low concentration of nH_s before the robot would interact with it. It should be stressed that these value are predetermined and it this experiment no learning takes place. Therefore outside of perception modulation from the hormones (see equation 4.3), the novelty of each object type remains constant.

The robot's perception of the perceived novelty is therefore determined by the following:

$$PerceivedNovelty_i = InitialNovelty_i + StressScaling_i \times nH_s + CurosityScaling_i \times nH_c \tag{4.3}$$

where $PerceivedNoveltyt_i$ is the final perceived novelty value of object type i, $StressScaling_i$ and $CurosityScaling_i$ is the effect that the current concentration of hormones nHs and nHc respectively have on the robot's perception of the novelty value (see table 4.1)

This perception of novelty is then used to modulate the robot's behavior by changing the previous formula 2.9 to the following:

$$WheelSpeed_i = \sum_{v,d} motivation_{v,d} \times directionScaling_{i,d} + PerceivedNovelty_d \tag{4.4}$$

Therefore if the robot detects an object with a high perceived novelty, given no other stimuli the robot will move toward it, and vice versa if the robot detects and object with a negative perceived novelty. Objects with negative perceived novelty are essentially considers objects which are two risky or dangerous to interact with at the given time.

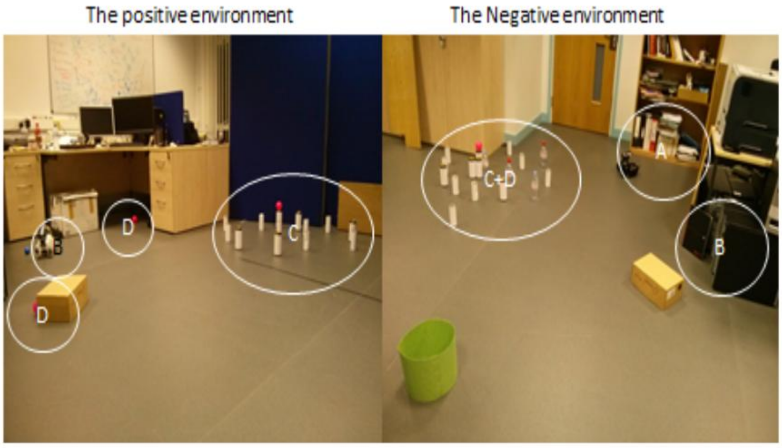
4.5.1 A new open environment

As part of the expansion of the robotic model we now looked to move away from the relatively small and sterile environments we had used in chapter 3. From now on all robotic experiments were carried out in our open lab space (around 7^2 meters excluding equipment/furniture) which can be seen later in figures 4.2 and 7.1. While as we will discuss the exact make up of this environment changes to suite the current experimental conditions, the robot is given general freedom to move and interact with any aspect of the environment. Not only should this provide the robot with greater opportunities to develop, but also give us an opportunity to look at how model works in a more "real-world" dynamic unstructured environment

The robot is only excluded from interacting with sensitive equipment which it could damage. In order to protect these areas from the robot simply plywood border are used.

4.5.2 The first part of the experiment

The first part consisted of placing two identical robots in two different environments: a "negative" or "stressful" environment and a much more "positive" environment, with one robot in each environment for a period of 20 minutes. Negative environments were designed to be more hazardous and less rewarding, for example requiring the robot to navigate smaller spaces. In contrast positive environments were designed to provide the robot with both adequate novelty and appropriate reward. Both of these environments took place within the previously described environment in section 4.5.1 the following changes made which can be seen below and also shown in figure as can be seen in figure 4.2 and discussed below.


The positive environment

The Negative environment

Fig. 4.2 The positive and negative environments. The positive environment is on the left and negative environment on the right with the differences discussed in section 4.5.2

- Firstly, the homogeneity of materials used was different, with more variation occurring in the negative environment (label A in figure 4.2). As IR Sensors naturally respond differently to different materials, increasing variation naturally leads to more fluctuations in sensor readings. Due to the previous described artificial HPA/HPG axis, these fluctuations are likely to lead to a stress response.
- Secondly, the increased use of objects made of textures and/or colors particularly those that are difficult to detect (I.e Highly reflective objects) the negative environment (label B in figure 4.2). Not only does this lead to fluctuations in sensor readings but also,

since the robot is unable to accurately detect distance, it also increases the likelihood of collisions as the robot is less able to accurately predict the distance.

- Thirdly, is a difference is the spacing between objects (label C in figure 4.2). In the negative environment, distance between different objects is small, increasing the potentially for stressful encounters. This is due to the smaller spaces between objects increasing the chance of collisions, over-stimulation (due to exposure to multiple objects at once) and becoming lost or trapped and therefore unable to sate other needs.
- Fourthly is the "reward" obtained for exploring the environment, which is greater in the positive environment (label D in figure 4.2). For example, in the positive environment the reward for completing the maze (See figure 4.2) is an easily accessible homeostatic resource. In contrast, while the resource is still present in the negative environment, the likelihood of finding and accessing the resources, and therefore of getting a reward for completing the maze, is smaller.

These environments where designed to lead to different developmental trajectories of each robot. Due to the epigenetic mechanism and the HPA-HPG axis we would expect that the robot placed in negative environment would develop a more withdrawn phenotype and due to stressful exposures would be less interested in exploring and interacting with sources of novelty. In contrast the robot placed in the positive environment should be expected to develop a more outgoing phenotype being more interesting in exploring novelty due to its past experiences.

In order to test this hypothesis after spending time in their respective environments the robots where then placed (individually) in two more environments in order to compare the effects on the original environments on the robot's behavior.

- Firstly a set of 10 single-robot runs, each of duration of 15 minutes, took place in relatively static environment, which consisted of our unmodified lab as preciously shown and discussed in section 4.5.1.
- Secondly a set of 10 runs of a duration of 5 minutes per run were under a Human-robot Interaction setup (H-R). In this second experiment, human-robot interaction took place in an empty environment. While challenges and potential stimuli to explore were still in abundance at the edges of the environment, the center was largely barren in order to increase the chances that the robot would focus on human interaction.

For the H-R experiment were reduced to 5 minutes since the experimenter could finely control exposure to stimuli, making unnecessary the longer duration of runs that was used in the static environment condition in order to ensure that the robot could fully explore the environment. While the sensory-driven robot model and classification network were never specifically designed or programed for human-robot interaction, we found that it would naturally lend itself to simple types of interaction.This is due to both the categories set out in table 4.1 and the physical workings of the robot's sensors.

For example, the robot generally found stroking interaction positive as it led to a release of the curiosity hormone, whereas sudden movement usually invoked a fearful or negative response. This was due to slow stroking normally being identified as the detection of a single object, where as rapid movement trigged multiple sensors and therefore was often identified as either a Large object or a Hole (see 4.1)

Of course the actual response of the robot to the interaction is also influenced by a combination of the robot's internal state and developmental history. A robot which has a withdrawn phenotype or a high concentrations of the stress hormone is likely find any type of interaction negative. In contrast a robot that has a more outgoing phenotype might "enjoy" the rapid stroking.

For both robots the H-R experiment followed a predetermined script of interaction.

4.5.3 Results of the first part of the experiment: The "static" Environment

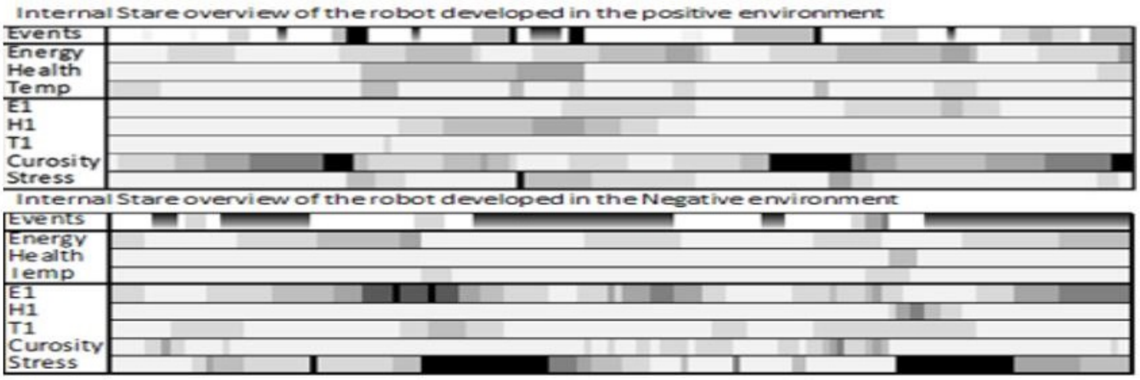


Fig. 4.3 The internal state of the two robots during the 15 minutes in the static environment, with the robot from the positive on top and robot from the negative environments below.Darker colors indicate increased stimuli, higher homeostatic deficits or increased hormones concentrations respectfully

As can be seen in figure 4.3, the interaction between the environment and epigenetic hormonal mechanisms made the robot have significantly different behaviors once developed and placed within the “neutral” testing environment.

In all cases, robots that had developed in the negative environment showed a very “withdrawn” behavior: the robot spent a significant portion of its time executing a behavior similar to wall following. If the robot found a corner or an enclosed area, it would remain stationary in this location until other internal needs (e.g., the need to replenish energy) became more prioritarian. The reason for stopping in these enclosed areas was likely to be the fact that they were perceived as the safest location – as walls, which, detected on multiple sides, in a stressed state would have positive valence. This behavior could perhaps be seen as the robots treating corners as nests, using them to reduce current stress levels.

The reason for this withdrawn behavior can be traced to the highly unregulated stress receptor as a result of the “stress” the robot was exposed to in the negative environment.

This up-regulated receptor had two significant effects on the robot’s behavior …

- Firstly, the up-regulated stress receptor meant that even minimal or basal levels of the stress hormone could lead to a significant stress response. This meant that all objects in table 4.1 which scales negatively with stress were almost always perceived as having a negative valence, an where therefore avoided

- Secondly, due to the robot’s high sensitivity to stress, the secretion of the curiosity hormone was almost fully suppressed most of the time.

As can be seen in figure 4, interaction with other areas of the environment was minimal due to the constant “stress” levels, which suppressed the HPG-Axis effectively, preventing the emergence of a ratio between curiosity and stress that would modulate the robot model into investigating novel objects.

In a few rare occasions that the robot did have a high enough level of curiosity hormone to facilitate and initiate interaction with novel objects, it quickly became over-stimulated and reverted to the previous withdrawn behavior. Stress responses to interactions with novel objects not only tended to be more prevalent in these robots but also were significantly more severe and lasted on average 60 percent longer. Stress responses and hypersensitivity to homeostatic deficits where also heightened in these robots. Essentially, this meant that the robot would look to maintain homeostatic deficits at a higher level and if they started to drop, the robot would quickly enter it withdrawn behavior.

The implication of this is that, once the robot found an area of the environment with access to both resources, it would tend to stay in that general region and never really explore for new opportunities.

In contrast, the robots that had developed in the positive environment showed a much more outgoing behavior, thoroughly exploring the entirety of the environment and interacting with a large range of the different novel objects.

The reason for the outgoing behavior in these robots was almost the exact opposite of the robots from the negative environment. Here, positive experience during development led to an up-regulation of the Curiosity receptor. Leading to the robot to be more attracted to novelty

While this outgoing behavior did lead to increased risks such as collisions or over-stimulation, which caused the two high stress moments that can be seen in figure 4.3, the robot recovered fairly quickly.

In addition the robot developed in the positive environment tended to have a greater tolerance to homeostatic deficits, which resulting in it spending more time exploring and interacting with the environment. This was due to both the increased attractiveness to different aspects of the environment and due to slightly down-regulated eH receptors in comparison to the negative robot. These down-regulated receptors likely occurred due to both the increased rewards and ease of navigation in the positive environment (see section 4.5.3)

4.5.4 Results of the second part of the experiment: Human-Robot interaction

In the second part of the experiment, dynamism was introduced by the presence of a human who interacted with the robots. Once again, the robots in the different runs had developed either in a positive or a negative environment. Due to the limited range of the robot's sensors, lack of learning and lack of explicit programing for Human to robotic studies, the range of "recognizable" interactions was relatively small. However, the robots where able to detect movement and respond to simple interaction.

As we could expect, the robot that had developed in a negative environment had a "timid demeanor" and tried to avoid any form of human interaction. However, gentle stroking motions along the IR-sensors could be used to initiate interaction by causing a rise in the concentration of the curiosity hormone. Interaction was primarily limited to this slow stroking as well as the robot exploring the human at its own pace.

Any sudden movements or overzealous stroking would quickly lead to over-stimulation of the robot and an attempt to withdraw due to secretion of the stress hormone. Even with an ideal levels of interaction, the hypersensitivity to homeostatic deficits meant this robot would only spend a maximum of around 30 seconds interacting before becoming more interested in procuring resources.

As we also expected, the robot that developed in a positive environment was much more tolerant of interaction with the human. Slow- to medium-speed stroking led to an initial positive response; after a period of interaction, faster stroking and sudden movements were tolerated and even sparked interest from the robot due to its rising curiosity hormone concentration.

To this extent, if an object such as a ball was thrown, the robot would go after it to investigate. Once the ball/object stopped and the robot had explored it as a normal novel object in the environment, the interest in the object would drop, often leading to the robot to return to the human in search of increased stimulation.

It is worth noting that the robot did not know who or what had thrown the object and returned to the human purely because s/he is a large moving object and therefore had a high positive valence). As a comparison, an object thrown at a robot from the negative environment almost always led to the robot's withdrawal.

4.5.5 Conclusion

In this chapter we have looked at the potential of introducing a hormone based curiosity seeking mechanism. This mechanism should ideally regulate the novelty seeking behavior of the robot in an appropriate manner given the robot's current internal state and past exposures. What was shown was that if the robot is developed in a negative environment it will suppress its curiosity seeking behavior, preferring instead to simply stay in a safe location while maximizing its homeostatic levels, providing a buffer to help protect itself from perceived environmental dangers.

It is only once this robot has achieved this buffer that it will begin to display curiosity seeking behavior. Even then this behavior is limited with the robot quickly become stressed and withdrawing to perceived safety. In contrast, a robot developed in a positive environment spent a large portion of its time interacting with and exploring its environment.

The robot's response to an action was also highly influenced by the way it was carried out. For example like in [23], the speed and duration of the stroking motion will have a significant impact on the robot's response. This included interaction with a human, for which the robot had not been programmed and that emerged as a consequence of the developmental history of the robot. Based upon these result it appears that the artificial HPA-HPG axis does indeed provide the desired adaptive novelty seeking behavior. For example a robot that is developed in the negative environment as would be expected is more cautious and less likely to take risk associated with interacting with sources of novelty. Obviously this is only first stage of testing the novelty mechanism, now we are happy that it does provide the desired result in a simple sensory driven robot we will now integrate this into our learning system in the next chapter in order to see if combined, the ENN and artificial HPA-HPG axis can result in an open ended novelty driven learning system

Chapter 5

A emergent Neural network for adaptive Learning: The ENN

With the artificial neurohormone demonstrating the ability to generate and regulate curiosity driven motivations and behavior, in this chapter we will introduce and explore the ability to integrate these into our new artificial neural network the Emergent Neural Network (ENN). We hypothesize that by having both components, namely curiosity and learning present and interacting with each other, a learning system will emerge that enables the robot to learn and develop in an appropriate manner given its current and past environmental conditions. Essentially we suggest that these hormones should lead to the creation of an internal environment which supported appropriate learning and behavior, i.e., a chemical soup could regulate neural development of an emergent neural network in an appropriate manner given the robot's environment. In this chapter, we will focus the discussion on introducing the ENN architecture. In the following chapter 6 we will look at testing the hypothesis that the hormones can appropriately regulate the networks development.

5.1 Introduction and principles of the Emergent Neural Network

The proposed ENN architecture discussed in this chapter consists of a novel design it does bear some minor conceptional similarities to other forms of networks and learning models,

- Like with GasNets both models utilize chemicals which diffuse and modulate the network. Specifically, GasNets which are a type of Dynamic Recurrent Neural Networks (DRNNs) [96] utilize different abstract models of diffusing gaseous neuromodulation into the network. This gaseous neuromodulation, is inspired by real nervous systems in which different chemicals that diffuse through the system are shown to modulate the system [62]. In Gasnets this modulation tends to occur in the transfer function between two nodes. Like with a GasNet, in the ENN chemicals (our artificial hormones) diffuse through the network (see section 5.2.2) where similar they modulate the firing of the node. However, unlike a traditional gasnet our chemical can also influence the signal sent between two nodes.

- Like with the theory Hebbian learning (see [128] for an overview), the ENN share a similar concept of "Cells that fire together, wire together". In this sense when two nodes fire within the same action loop, the synaptic connection between them will strengthen or if is their is no connection to begin with, one may be made. However what may be different to atypical Hebbian learning implementation is that the synaptic connection between nodes will also reduce or even break if the nodes are not regularly active in the same time period. Additionally as will be seen later on, apart from similar concepts, the actual implementation between the ENN and typical hebbian learning have little similarity (see section 5.2).

This network essentially consists of the following characteristics

- The network has three distinctive layers consisting of an input, hidden and output layer. The Input and output layers possess a fixed number of nodes.

- The different layers all have slightly different rules and components and will be set out in their relevant sections.

- The input layers which consists of one node per “sense”, feeds in data from the robot’s different sensors.

- The output layer contains nodes which represent the robot’s homeostatic variable and has one node per variable

- The hidden layers starts empty with nodes created as a function of the robot’s interaction with different aspects of its environment.

- Nodes posses a synaptic plasticity which is used by the robot to determine the novelty of a stimuli.

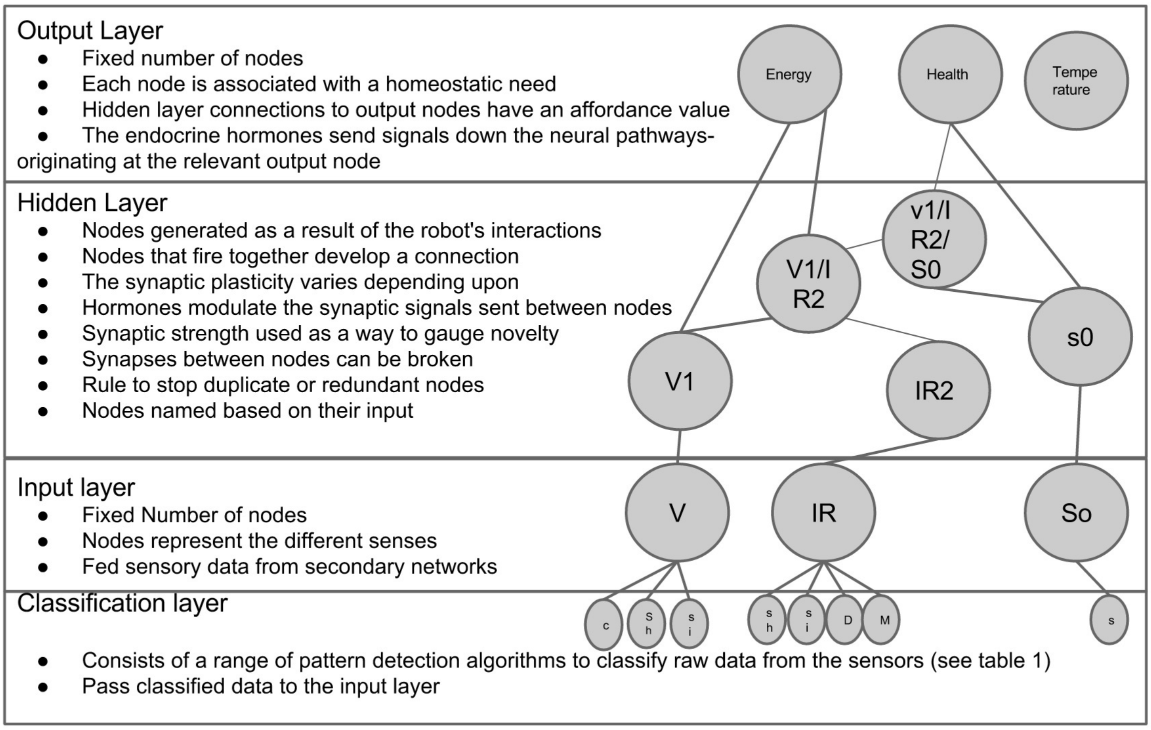
- Connections between nodes is made as a result of firing at the same time bearing some similarities to hibernian learning
- The robot's internal hormone concentration play a role in regulating the network
- The output from the emergent neural network directly modulates the wheels of the robot. It is through this modulation that different unplanned behaviour emerge.

5.2 Design of the Emergent Neural Network

The emergent neural network consists of a novel design in which nodes are created as a function of the robot's interactions and exposures to different environmental stimuli. This emergent neural network, of which an example can be seen in figure 5.1 is designed to allow the robot to learn the affordance of different aspects of its environment. Here the term affordance is used in the context of the robot learning the potentialities of an action or interaction with different aspects of its environment in relation to its current internal state. Since the internal state of the robot presented here is dependent upon, and made up of the three homeostatic variables (see section 2.2), the affordances learned by the robot will be in relation to the ability of actions to affect these said variables. For example, a potential action involving the energy resource, will likely have an affordance associated with energy recovery. By being able to learn these affordance, the robot no long relies upon the values in table 4.1. Rather all objects will start with an affordance of 0. Through the robot's interactions it will then be able to determine thee effect that the object has on the robot's behavior.

At this stage it is important to highlight two aspects of the robotic model:

- Firstly all behaviors that will be discussed emerge as a result of the development and modulation of the neural network, simply put there are no pre-designed behaviors or internal states. While certain hormones may encourage the robot to act in a predetermined way, (i.e the energy hormone, $E1$, encourages the robot to move towards a source of energy), the final behavior that emerges will reflect all intrinsic motivations and external stimuli. The greater the strength of a motivation or stimuli the greater influence it will have on the robot's behavior.
- Secondly the development of both the neural network and affordances are based upon the robot's perceptions and interactions, therefore robots with different morphological designs or placed in different environmental conditions will likely develop in different ways.



Output Layer
- Fixed number of nodes
- Each node is associated with a homeostatic need
- Hidden layer connections to output nodes have an affordance value
- The endocrine hormones send signals down the neural pathways- originating at the relevant output node

Hidden Layer
- Nodes generated as a result of the robot's interactions
- Nodes that fire together develop a connection
- The synaptic plasticity varies depending upon
- Hormones modulate the synaptic signals sent between nodes
- Synaptic strength used as a way to gauge novelty
- Synapses between nodes can be broken
- Rule to stop duplicate or redundant nodes
- Nodes named based on their input

Input layer
- Fixed Number of nodes
- Nodes represent the different senses
- Fed sensory data from secondary networks

Classification layer
- Consists of a range of pattern detection algorithms to classify raw data from the sensors (see table 1)
- Pass classified data to the input layer

Figure 1. An example of the basic structure of the neural network.

5.2.1 Basics of the emergent neural network

The emergent neural network consists of a three layer network design shown in figure 5.1. The first layer of the network consists of an input layer which is fed sensory data from a range of different classification networks. The second layer is the hidden layer in which nodes emerge as a function of the robot's interactions and environmental exposures. This layer is responsible for recognizing different aspects of the environment and assigning an appropriate affordance based on the robot's past experiences. The final layer is the output layer which simply sums the detected affordances.

5.2.2 Classification and input nodes

The input layer consist of three fixed nodes each representing one of the robot's different sensory modalities. These modalities are vision, IR and sound, and receive input data from different pre-processing classification algorithms shown in table 5.1 and figure 5.2. These three input nodes are queite different to conventional neurons found in other networks. Unlike conventional networks, these nodes will fire differently depending on which classification network is currently feeding input, with each node in the input layer associated with a specific

Table 1. The different sensory modalities and their implementation and design

	Vision			IR				Sound
Sense	Color	Shape	Size	Shape	Size	Distance	Movement	Volume
Sensor	Webcam	Webcam	Webcam	IR	IR	IR	IR	Webcam
Algorithm	OpenCV HSV	OpenCV Contour	OpenCV Contour	Pattern detection	Pattern detection	IR Value	Compare IR value	Sound
	White=0 Black=1 Green=2 Red=3 Yellow=4 Blue=5 Unknown=9	Circle=0 Square =1 Rectangle=2 Triangle=3 Crescent=4 Unknown=9	Small=0 Medium=1 Large=2 Unknown=9	Flat =0 Curved =1 Corner=2 Hole =3 Unknown=9	Small=0 Medium=1 Large=2 Unknown=9	Close=0 Medium=1 Far=2 Unknown=9	None=0 Small=1 Medium=2 Large=4 Unknown=9	Silent=0 Quiet=1 Medium=2 Loud=3

fixed group of classification networks (see Table 1). For example, the node representing the vision modality is associated with classification networks that detect Color, Shape and Size. These input layer nodes work as follows.

For each sensory modality, the output from each of the pre-processing classification networks (shown in table 5.1) consists of a 4-digit input pattern that feeds into the appropriate node in the input layer. The four digits provide information about the sensory modality used, the type of stimulus, the position of the stimulus with respect to the body co-ordinates of the robot, and the number of times that the stimulus has been detected in that sampling point. The number of pre-processing classification networks associated with each node of the input layer depends on the modality of the latter – three for vision, four for IR and one for sound (see table 5.1). For each input pattern received, each node in the input layer will either strengthen the connections with a node in the hidden layer corresponding to that input pattern, if a node has already been associated with it, or create a new node if the pattern is classified as novel. In any one time frame a node in the input layer can receive multiple inputs from each pre-processing classification network, and thus it can potentially create multiple new nodes in the hidden layer.

As an example, when perceiving the face depicted in figure 5.3, the vision node in the input layer would receive an input from the shape pre-processing classifier consisting of the four digits 1 (indicating the vision modality), 0 (representing a circle), 2 (if the face was directly ahead), and 3 (for the three circles: two eye circles plus the larger enclosing circle). It would additionally receive an input of 1, 4, 2, 1 (indicating, respectively, vision, crescent, ahead, one instance).

5.2.3 Hidden layer

The second layer of the ENN is the hidden layer, which receives data from the input layer and sends data to the output layer. This layer initially starts empty, and nodes are created as a

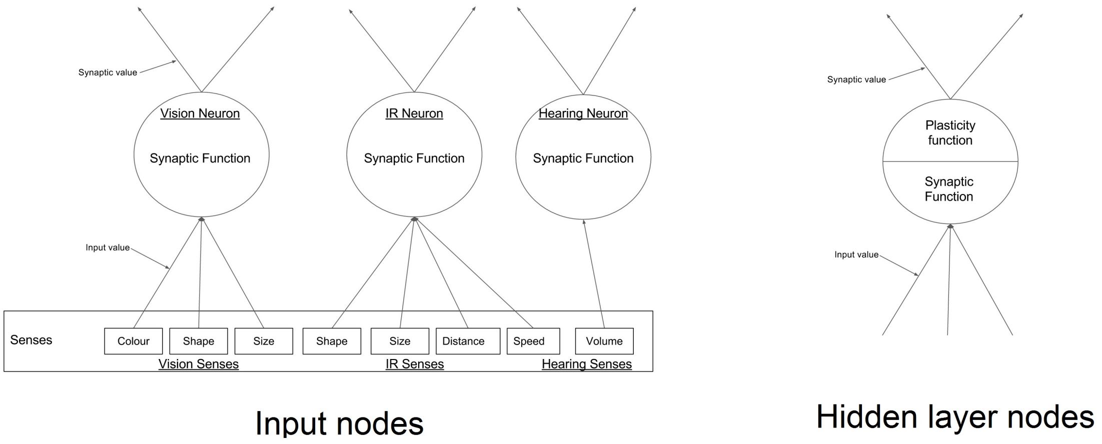


Figure 2. A generic example of the types of nodes in the ENN, the three input nodes, each representing one of the robot's senses can been seen on the left, and a hidden layer node on the right

function of the robot's exposure to different stimuli. Creation of nodes takes place under two circumstances:

1. When an input node fires but has no synaptic connection to a relevant node (as described in the section above), or
2. When two or more hidden layer nodes fire at the same time.

When a new node is created in the hidden layer, in addition to being connected to the relevant nodes that led to its creation (which provide the input), it is also fully connected to the nodes of the output layer. However, all these different connections can disappear as the network continues to develop. When a synaptic connection between two nodes is created, it is given a strength of 0.5. The connection strength is then updated as the robot interacts with its environment, using a sigmoid function, as seen in equation 5.1.

$$sP_{ij} = \alpha e^{\beta e^{x_{ij}e}} \tag{5.1}$$

where α and β are constants, and x_{ij} is the sum of times the nodes i and j have fired together, minus the number of times that they have not fired together, within a range of $-10000 < x_n < 10000$. A negative value of x_{ij} thus means that, more often than not, the nodes have not fired together.

Equation 5.1 results in a synaptic connection with strength in the range $0 < sP_{ij} < 1$. Due to the sigmoid nature of the function, the closer the synaptic strength gets to either end of this range, the lower the rate of change, or plasticity, of the connection. The synaptic strength of a connection between nodes plays a number of roles in this ENN, as will be discussed shortly. One of the most important roles is simply determining if a connection exists between nodes. This is achieved as follows:

- If a synaptic connection between two nodes exists but the synaptic strength drops below 0.05, then the connection is broken.
- If a synaptic connection doesn't exist but synaptic strength would be above 0.1 if it existed, then a connection is made.

In addition, when a synaptic connection is made from a new node to a node in the output layer, this connection is assigned an affordance – the potential to recover a homeostatic variable. Initially, this affordance is set at the change detected in the related output node's homeostatic variable. For instance, a new synaptic connection to the energy output node during a loop when the robot gained 2 units of energy will result in that connection receiving an initial affordance value of 2. This affordance assigned to the synaptic connection then changes as the robot continues to interact with that particular aspect of the environment, as follows:

$$\Delta Affordance_{v,i,j} = Affordance_{v,i,j} \times sP_{ij} + HomeostaticChange_{v} \times (1 - sP_{i,j}) \tag{5.2}$$

where $HomeostaticChange_v$ is simply the change in a homeostatic variable v in the current action loop compared to the previous one.

In order for a node i in the hidden layer to fire, it must receive a total input that is greater or equal to its total number of synaptic inputs, thus:

$$output_{i,d} = \begin{cases} 1 & \text{if } input_{i,d} >= sC_i \\ 0 & \text{otherwise} \end{cases} \tag{5.3}$$

where sC_i is number of input synaptic connections for node i, and d $(0 \leq d < 8)$ is the direction of the detected stimulus with respect to 8 equally spaced body co-ordinates of the robot, the third digit of the input pattern discussed in section 5.2.2. Using this system, 0 represents the body co-ordinate directly behind the robot, then going clockwise each subsequent value represents the next co-ordinate. For example 4 is represent co-ordinate directly in front of the robot.

As shown in the previous equation, if the firing threshold is reached, the nodes fire with a value of 1; however, the synaptic function $sF_{i,j,d}$, or output of the hidden layer node i , is then modified depending on the outgoing connection of the node and the directional origin of the stimuli d. If a hidden layer node i is connected to another hidden layer node j, the synaptic function is:

$$sF_{i,d} = output_{i,d} \times nHmodulation_i \times \sum_v eHModulation_{v,i} \tag{5.4}$$

where $eHModulation_{v,i}$ is the concentration of the endocrine hormones eH modulating node "i" (see equations 5.6 and 5.7) and $nHmodulation_i$, is the combined strength of the modulation from th neuro-hormones stress and curiosity nH (see equation 5.8). The roles of hormones in the ENN is discussed in greater detail in section 5.2.4.

If the node is connected to an output node then the synaptic function is given by:

$$sF_{i,j,v,d} = nHmodulation_i \times output_{i,d} \times Affordance_{v,i,j} \times eHModulation_{v,i} \tag{5.5}$$

A basic example of how the ENN works and allows the robot to identify objects and stimuli can be seen in figure 5.3, which shows how the robot perceives a simple face. Here, the robot is able to identify the face by the presence of the key characteristics of a large circle, 2 small circles and a crescent. However, the robot cannot detect spatial arrangements, and therefore, as long as the features are close enough, they will be identified as the same object. The characteristics used by the robot to identify objects depend on its past learning. A relatively new robot, for instance, may identify all pictures as being the same, since they all posses a circular shape, In contrast, a robot with greater environmental exposure, such as the one used in this example, will have more specific criteria.

5.2.4 Hormones in the ENN

As shown in equations 5.4 and 5.5, different hormone concentrations modulate the synaptic functions of the ENN. As previously discussed in section 5.2.3, these hormones modulate the nodes within the ENN as a function of the internal state of the robot. In the case of the *eH* hormones, the strength of the modulation is dependent upon the hormones receptors sensitivity and the connections of the nodes. The modulating effect of each of the three *eH* hormones starts at their the associated output node. For example modulation from the E1

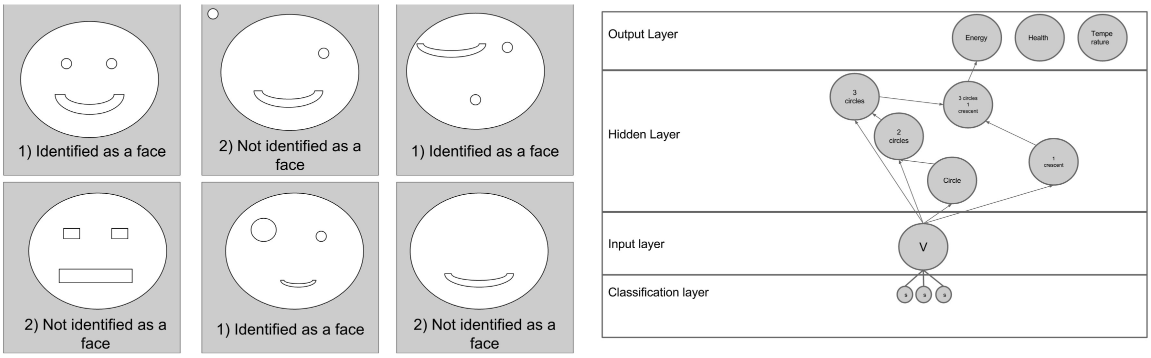


Figure 3. To provide an example of how the robot perceives its environment, we have shown the robot a simple picture of a face, seen in image 1) on the left. A simple example on how the ENN may develop in relation to this picture is then shown on the right. The robot is shown the 6 pictures on the left to see which ones are identified as being the same. In this particular example, the robot has learned to identify the original by the presence of a large circle, 2 smaller circles and a half crescent; hence, samples 1,3,5 on the left are all considered by the network to be the same face.

hormone starts at hidden layer nodes directly connected the Energy output node (node E in the output layer of figure 5.4). The hormone modulation (which is still determined by the previous formula 2.3) the effects subsequent nodes depending on the synaptic plasticity. If a hidden layer node has no direct or indirect synaptic connection to a specific output node then it will not be modulated by the associated hormone concentration i.e., if node *i* has no connection to output node E, it will not be effected by E1.

For a node *i* directly connected to the output layer, the modulation by the eH_v is given by:

$$eHModulation_{v,i} = eHConcentration_v \times senS_v \tag{5.6}$$

where $eHConcentration_v$ is the concentration of the $eHhormone$ eH_v, $senS_v$ is the sensitivity to $eHConcentration_v$ (See equation 2.10).

However, for nodes not directly connected to an output node, the modulation from the eH_v may become weaker depending upon synaptic strength between nodes, as shown in equation 5.7, resulting in a larger modulation of the nodes closer to the output layer and/or with stronger synaptic connections to it. Using hormonal modulation in this manner promotes the activation of nodes that have a higher synaptic strength, and hence promotes behaviors that, in past interactions, have led to better homeostatic balance.

$$eHModulation_{v,i} = \sum_{j \in O(i)} \frac{eHModulation_{v,j} \times sP_{i,j}}{noI_j} \tag{5.7}$$

where $0(i)$ is the set of output nodes from node i, i.e the set of nodes that are connected to output of i, $eHModulation_{vi}$ is the strength of the modulation in the current node, dependent on the sum of the signal passed down from connecting nodes $eHModulation_{vj}$, and noI_i the number of input connections of node i.

In contrast to the *eH* hormones, the *nH* hormones surround the ENN, affecting all nodes equally. The *nH* behave differently as their role is to either promote or suppress novelty-seeking behavior. This is caused by the combined effect of the curiosity and stress hormones. The curiosity hormone increases the activation of nodes with a low synaptic strength, and suppresses nodes with a high synaptic strength. Conversely, the stress hormone increases the activation of nodes with a high synaptic strength, and suppresses nodes with a low synaptic strength, as in equations 5.4 and 5.5. Therefore, the robot is using the synaptic strength as a way of assessing the novelty value of an object or aspect of the environment, since a high synaptic strength only happens if an object behaves as expected each time the robot interacts with it. This can be seen below in equation 5.8

$$nHmodulation = sP_{ij} \times nHConcentration_s \times senS_s + (1 - sP_{ij}) \times nHConcentration_c \times senS_c \tag{5.8}$$

where $nHConcentration_s$ is the concentration of the stress hormone (s) and $senS_s$ is the receptor's sensitivity to the stress hormone. $nHConcentration_c$ is the concentration of the curiosity hormone (c) and $senS_c$ the receptor's sensitivity to the curiosity hormone.

5.2.5 Output Layer

The final layer of the ENN is the output layer, which consist of a fixed number of nodes equal to the total number of survival-related homeostatic needs. Each output node simply sums up the total input from the hidden layer in order to calculate the affordance of moving in a certain direction.

$$output_{v,d} = \sum_i sF_{i,v,d} \tag{5.9}$$

The output of the ENN then feeds directly into (and modulates) the robot's actuators – in this case the wheels:

$$WheelSpeed_i = \sum_{v,d} output_{v,d} \times set_{i,d} \tag{5.10}$$

where $WheelSpeed_i$ is the speed of the left (0) or right wheel (1), set_{id} are constant vectors equal to $(-10,-10,-5,-3,1,3,5,10)$ if $i = 0$, or $(-10,10,5,3,1,-3,-5,-10)$ if $i = 1$. This means that if a single stimulus originating from the left side of the robot is detected, the robot's left wheel moves at a speed of $-5 \times theoutput$ and the right wheel moves at a speed of $5 \times theoutput$. Therefore, a positive output will result in the robot turning towards the stimuli and a negative away from it.

To summarize, the causal chain that leads to internal or external stimuli promoting different behaviours is as follows

1. As homeostatic deficits occur they lead to the release of the associated endocrine hormone eH, (see equation 2.3).
2. Internal and external stimuli lead to the release of the neuro-hormones, with curiosity being secreted in relation to perceived positive stimulation and stress in relation to negative stimulation (see equations 4.2 and 4.1).
3. The robot's sensitivity to each hormones is dependent upon its historic exposure to it (see equation 2.10).
4. The artificial hormones modulate the synaptic function of the hidden layer nodes (see equation 5.5) depending upon the nodes position in the network (see equations 5.6 and 5.7).
5. The output nodes sum up the synaptic function of connected neural pathway (see equation 5.9), with the value dependent upon past outcomes associated with pathways activation (see equations 5.1 and 5.2).
6. The output from the network then effects the behavior (wheel speed), by promoting or suppressing the desire to move in a certain direction at a certain speed. The larger the output, the greater modulating effect it will have on behavior (see equation5.10).

The complete model including both the HPA-HPG axis and ENN components can now be seen in figure 5.4

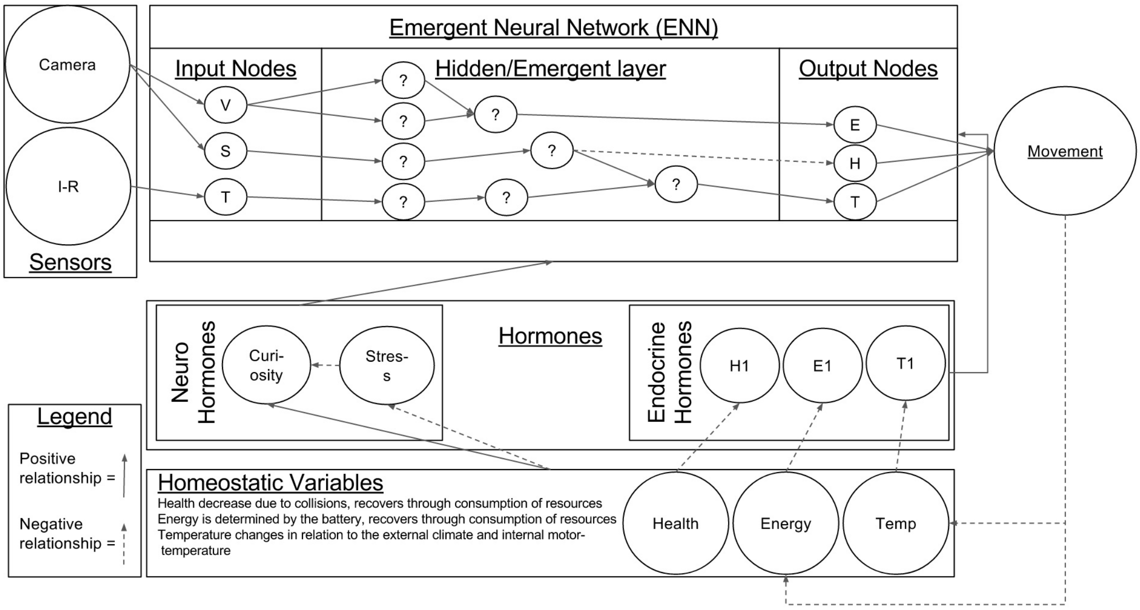


Fig. 5.4 The final robotic architecture

5.3 Additional changes to the model: Moving away from virtual homoeostatic variables

At the same time as the addition of both the ENN and HPA-HPG axis, interdependently we also began looking at moving away from the virtual homeostatic variables described in section 2.2. The plan instead was to implement homeostatic variables based on the actual physical state of the robot.

5.3.1 Energy

To this extent the energy variable was changed to the following.

- Energy is now linked directly to the robot's battery which has a total charge of 3500 mAh and decreases at an average of around 15 mAH per minute, although the exact amount varies as a function of the robot's motor usage.
- As this battery would give the robot nearly 4 hours of running time without needing to recharge, which is considerable longer than are experiments and would negate the need to recharge, the robot is programmed to only sense a maximum charge of up to 75 mAH (around 5 minutes of running time). Essential creating a virtual battery

Fig. 5.5 The Virtual battery

- In order to recharge (sense more of the battery), the robot needs to find the same blue balls as before. Upon finding these blue balls the maximum charge sensed would return to 75 mAH

5.3.2 Temperature

In addition to Energy we did look at also changing temperature from a virtual measure to incorporating the robot's real physical internal temperature. This was achieved using an internal heat monitors, with the system used in [80]. However, we have since found that using the robot's real temperature was idle for the length of experiments carried out here (due to length of time taken for the robot to heat up and more importantly cool down) For this reason we continue to use the Virtual temperature per equation 2.2.

5.3.3 Health

For obvious reasons (we do not want to physical damage the robot!) we made no attempts to try to link health to the robot's physical condition. Therefore health has remained the same as equation 2.1.

Chapter 6

Epigenetics and situated learning through the ENN: The final model

In chapter 4, we investigated and demonstrated how an artificial epigenetic neuro-modulatory mechanism inspired by the roles of the steroid hormones Testosterone and Cortisol can regulate the behavior of an autonomous robot in relation to both its current and past internal and external environments. The implementation of this artificial hormone system works through the generation of internal drives and the modulation of perception (see section 4.3) leading to appropriate motivations and subsequently appropriate behaviors. For example, a robot which had a low wellbeing, perhaps due to low homeostatic variables or over-stimulation would be less inclined to engage in risk taking or interacting with novelty, instead preferring low risk or safety (see section 4.5). On the other hand a robot which has a high wellbeing would be more likely to take risks and explore novelty. Further, due to the epigenetic nature of the system, the robot's past interactions and histories influence its current perception indirectly through the regulation of hormone receptors (see equation 2.10). This means that even with the same homeostatic levels, A robot which has consistently been exposed stress or otherwise unpleasant experiences will be less inclined to engage in risk taking behaviors than say a robot which has had more positive experiences (see section 4.5).

In chapter 5 we introduced the ENN neural network which was designed to allow the robot to learn the affordance of different aspects of the environment through its interactions with them. In this chapter we bring together the neruo-modulation system and ENN into a signal final model. Specifically we look out these two systems can be integrated to create a flexible adaptive learning system. We suggest that with both systems present (see figure 5.4) the

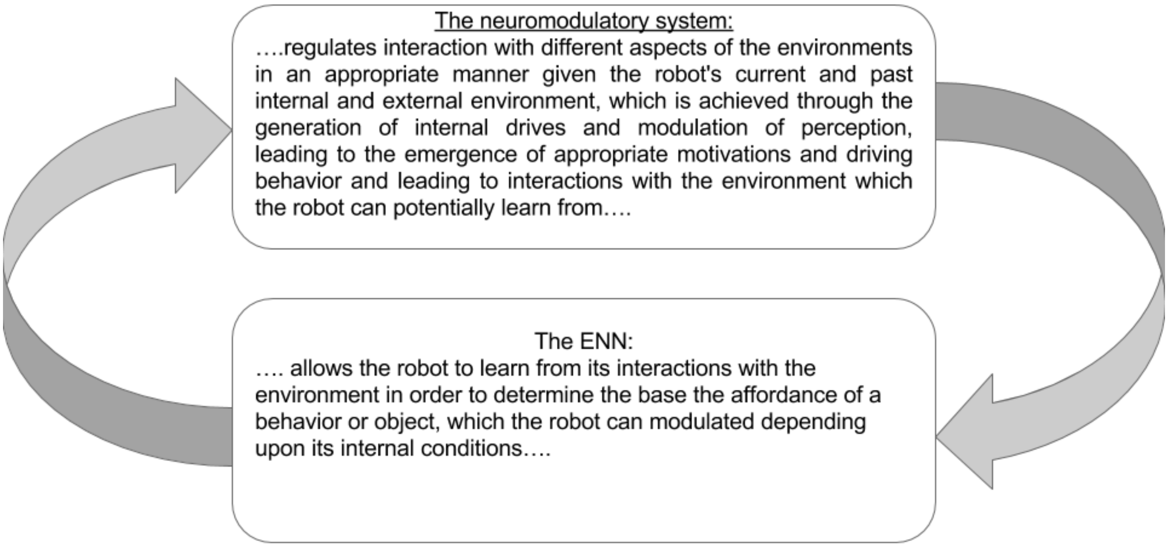


Fig. 6.1 The casual relationship between the ENN and the neuromodulatory system

robots can learn in a way which is appropriate for it given environment due to the way they influence each other as shown in figure 6.1).

With the neuro-modulatory mechanism (chapter 4) and ENN (see chapter 5) set out, in this chapter we can test our original hypothesis that through hormone secretion, our epigenetic mechanism can create and regulate a internal environment (through the creation of diverse chemical soups) which supports appropriate neural development given the current environmental conditions. In order to test this hypothesis a robot was developed under varying environmental conditions, the outcomes of which we report in this chapter.

6.1 Operating the epigenetic model in a dynamic open environment

In order to test if the combination of the ENN and epigenetic neuro-modulatory system do work together to create an adaptive learning system an experimental setup was designed which consisted of 3 different environmental conditions with a different robot developed in each environment. In these environments experimental variation was introduced through the use of a physical person called the "caregiver". As discussed in section 6.2 the caregiver's behavior will vary in each experiment resulting in contrasting experiences for each robot.

Excluding the variations of the caregiver's behavior, all other environmental factors remained consistent across the experiments. If the proposed combination of the ENN and epigenetic neuro-modulatory system do lead to an adaptive learning experience for robot then we can expect the following.

1. That each robot will develop different behaviors to suite their current environmental conditions different conditions.
2. The varying behaviors will results in the robot having different learning experiences and approaches which are more suitable and effective for their current environmental conditions

In the following section 6.2 we will describe the experimental paradigm designed to test the above.

6.2 Experimental setup

Like with the past experiments we once again use the Kola II robot which is placed within the same open lab environment with only delicate areas, which could damage or be damaged by the robot, blocked off. In order to allow the robot to maintain its internal variables, 4 of the homeostatic resources (2 of each type) where also placed in lab, with one in each corner. In total 3 experiments of 60 minutes were run, with the only variations between each experiment being the behavior of the caregiver who played a different role in each experiment, these roles consisted of a neutral caregiver (see section 6.2.1) , a overbearing caregiver (see section 6.2.2) and a caregiver (see section 6.2.3). The robot's motivations were shown on a terminal in order to allow the caregiver to monitor the its current internal state and decide how to interact with it, depending upon their current role.

6.2.1 Neutral caregiver

In this first environment the robot was placed in the environment with the neutral caregiver who largely left it alone. The neutral caregiver would make no effort to actively engage with the robot. Rather in this experiment the caregiver would go about a "normal" routine in the lab as if the robot was not present. This routine consisted of the caregiver engaging in predetermined mundane behaviors at a set time and independently of robot, such as walking across the room to get an object from a cupboard at the 5 minute mark. The caregiver

Table 6.1 The interactions available to the caregiver to sate different desires of the robot

Interaction	Effect on the robot
"Feed" energy resource	Recovers energy
"Feed" health resource	Recovers health
slow stroking interactions/Remove novel object if applicable [1]	Reduces Stress
Bring a novel object to the robot	Promotes curiosity
Allow robot to stay stationary	Reduces temperature

and robot can be thought of as co-inhibitors of this environment, rather than having a dyad relationship. Interaction between the two would only occur if the robot and caregiver paths crossed independently or if the robot actively moved towards the caregiver. In cases where interactions occurred the caregiver would simply try to move around the robot. Therefore this environment offered the robot the most freedom in exploring and learning about it surroundings without interruption or distraction from the caregiver.

6.2.2 Overbearing caregiver

In the second of the three experiments the robot was placed in the environment along with an "overbearing" caregiver. In this scenario the overbearing caregiver will tailor their behavior to suit the robot's needs and seek to immediately sate any desires of the robot. This means for example that if the robot has a desire for energy the caregiver will immediately seek out an energy resource and bring it to the robot. A critical element to emphasis here is that all interaction between the caregiver and the robot are driven by the robot's desires. The caregiver here will never attempt to encourage any new desires or behaviors from the robot regardless of any perceived potential long or short term benefits. For example in the unlikely scenario that the robot was overheating but showed no desire to reduce its internal temperature then the caregiver would not interfere or attempt to cool the robot.

In this scenario the caregiver was able to perform the actions in table 6.1 in order to sate the robot's desires. In order to be able to monitor the robot's needs, the current highest motivation was displayed on an external monitor. If the robot had no current motivations, the caregiver would remain in the immediate vicinity though would not directly interact with the robot.

Table 6.2 The hostile behaviors available to the caregiver with a rough estimate of their effect on the robot, the actual effect of the interaction will vary depending upon the intensity and duration

Action available to the caregiver	Stress impact	Health impact
Feeding robot an unwanted resource	Low	N/A
Removing a resource robot is consuming	Low-High depending upon the need for that resouce	N/A
Rapid stroking/movement	Low-high depenging upon health level	low
Chasing the robot	Medium	N/A
Suppressing the robot (appearing)	Medium	N/A
Hitting the robot	Medium-High depending upon force	low-High depending upon force
Picking the robot up	High	Medium

6.2.3 Hostile caregiver

In this experiment a caregiver who was “hostile/unsuitable” for the robot was placed in the environments. Here, this caregiver would use behaviors which could be perceived as aggressive or stress inducing. These behaviors included chasing, rapid movements, excessive and or aggressive physical contact or simple unsuitable behaviors such as trying to feed the robot the wrong homeostatic resource (see table 6.2 for the full list of actions available to the caregiver). The caregiver’s actions used during this experiment where based on the robot’s current internal homeostatic conditions to ensure we avoided the robot “dying” due to excessive physical contact (health loss). Less intense interactions were used when the robot had a low wellbeing and more intense action used when the robot had a high wellbeing. So while the environment is both hostile and challenging the robot can survive.

6.3 Overview of runs

In the following section we will provide a brief overview of the outcome of the three different runs, however firstly will briefly highlight the type of interactions that could occur between the robot and the caregiver

6.3.1 Five observable behaviors which emerged during interaction

As part of the qualitative analysis of the robot’s observable behaviors that emerged due to interactions with the caregiver, we will often refers to 5 categories of behaviors consisting of Interest, Aggression, Aversion, Fear and Ignore. Interactions with the caregiver, like all behaviors discussed in this thesis emerge due to the modulation of the architecture and are not predetermined or preprogrammed. These 5 types of interaction that emerge will therefore have variations depending upon the robot and it current state.

- Interest is any behaviors which sees the robot move towards the caregiver excluding those classified as aggression. The underlying basis for these behaviors are generally a combination of curiosity driven (i.e the caregiver is perceived as being novel) or consummation driven (i.e the robot needs energy and believes the caregiver can recharge it) motivations.

- Aggression is any behaviors where the robot deliberately (i.e., non-accidental collisions, determine by looking at the robot's internal state) tries to push or bump the caregiver. Two different underlying bases were found for this behavior.

    1. The robot is trying to get to an object/place (i.e a resource) and the caregiver is in the way, rather than going around the robot tries to push past.

    2. The robot perceives the caregiver as a threat (due to past interactions) and "attacks" them to drive them off to prevent future homeostatic loss. This behavior is mostly observable when the robot has high homeostatic levels and/or has found attacking to be successful in the past.

- Aversion is any behavior excluding those classified as fear, which sees the robot attempt to avoid the caregiver by moving in a different direction. The basis for this behavior tends to be either or a combination of low homeostatic levels, particularly low curiosity leading to the robot trying to minimize risk, or perceiving the caregiver as a threat due to past interactions or uncertainty.

- Fear is a stronger form of aversion. Specifically Fearful behavior will see the robot move away from the caregiver (generally the opposite direction) with limited regards for other environmental stimuli. For example during behaviors classified as fear the robot often has increased collisions as the drive to move away from the caregiver overrides the drive to avoid an object. The basis for this behavior is simply a stronger version of aversion.

- Ignore is simply any behavior where the robot can detect the caregiver but the robot makes no attempt to interact. This will occur for two main reasons. Either a larger stimuli draws the robot's attention i.e the robot see a resources it needs, and/or given the robot's current internal state and past interactions with the caregiver they simply do not provide a large enough stimuli to affect the robot's behavior one way or another.

6.3.2 Pre-training

Before the robot was placed in the varying environment it first spent 10 minutes in a training environment. During this stage the robot was taught the basics of its environment such as what the different homeostatic resources looked like and basics of it own body such as collision resulting in health loss or movement leading to temperature gains. This learning took place in a relatively quiet and sterile environment so as to allow the robot to focus on the critical elements. Rather than do this three times, one for each in each environment, the training period was conducted once with a single robot. The ENN that had developed during this training was then saved at that point and loaded into the robot before the experiment. This ensured consistency and that each time the robot was placed into a new environment it did so with the same level of knowledge.

6.3.3 Neutral caregiver

In the first of the three experiments the robot was placed into the environment alongside the independent caregiver. Under these condition both the caregiver and robot were independent co-inhibitors. During this period the caregiver would not try to initiate any interactions with the robot rather they would simply move about the lab as they needed. Any interaction would need to be either initiated by the robot which would do so depending upon its current internal state and past interaction, or alternatively occur if both the robot and caregiver moved into the same area independently. In total during this experiment the caregiver was present in the environment for a period of 25 minutes and detected by the robot for just under 20 minutes. Over the course of this experiment the robot's behavior suggested it went through two developmental phases as indicated by the change in hormone levels in figure 6.2. These phases roughly correspond to the first 15 minutes of the run where the robot's behavior was focused primarily on the maintenance of it survival related homeostatic needs as it adapted to it's environment, and then a second phase for the remainder of the experiment where robot showed greater interest in novelty, exploration and learning.

The first 15 minutes: Adapting to the basics of the environment

At the start of this experiment, for the first 15 minutes the robot's behavior resembled a random walk with the robot internal drives leading it to explore the environment in order to

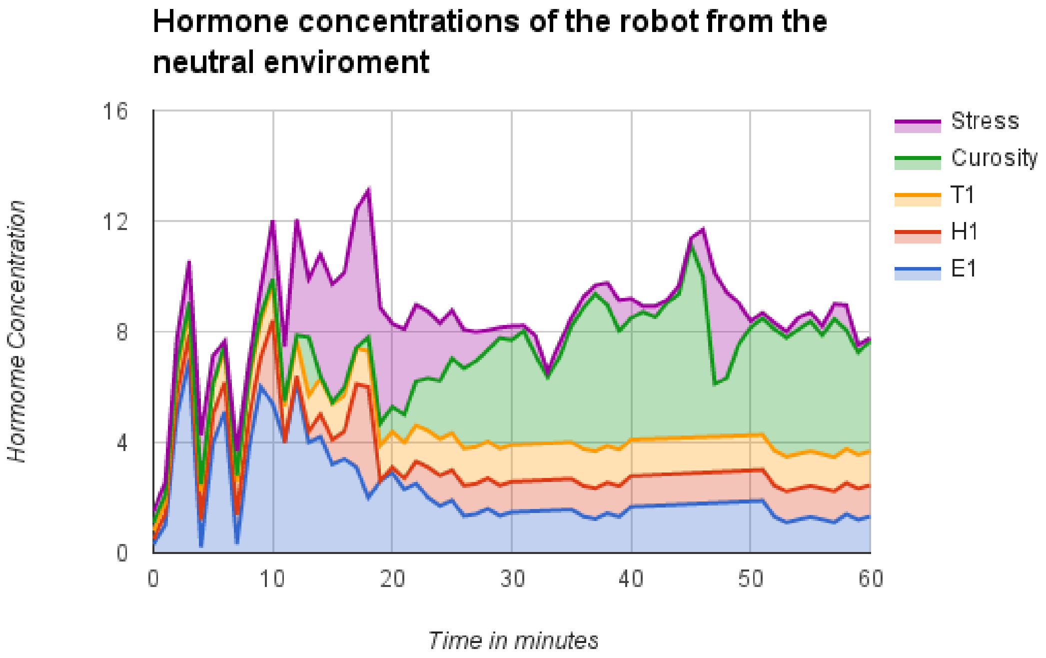


Fig. 6.2 The robot from the neutral environments hormone concentrations.

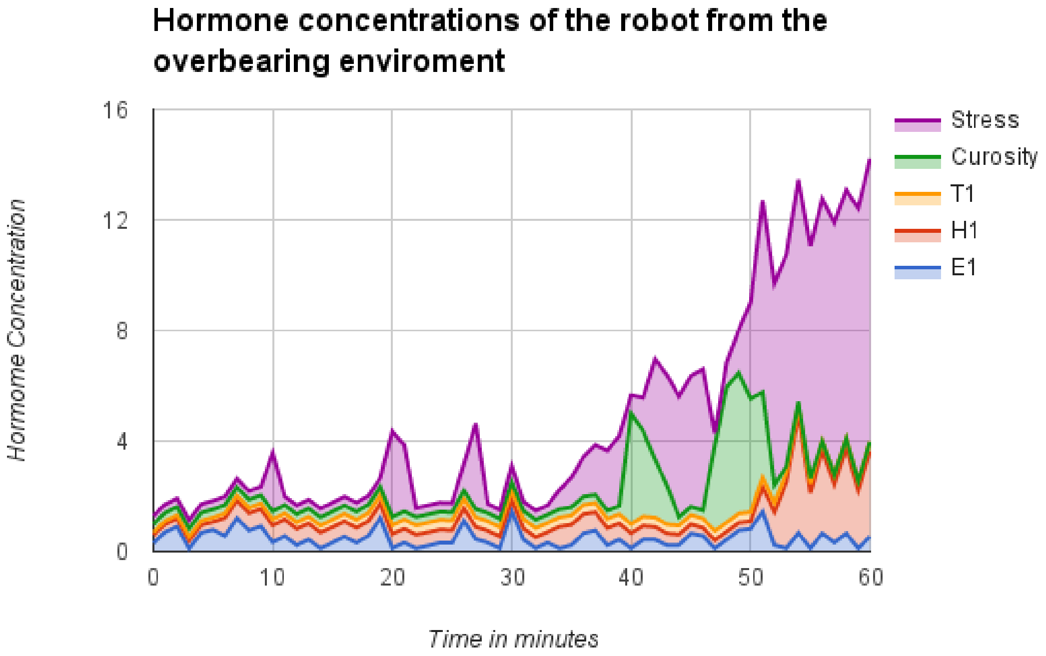


Fig. 6.3 The robot from the overbearing environments hormone concentrations.

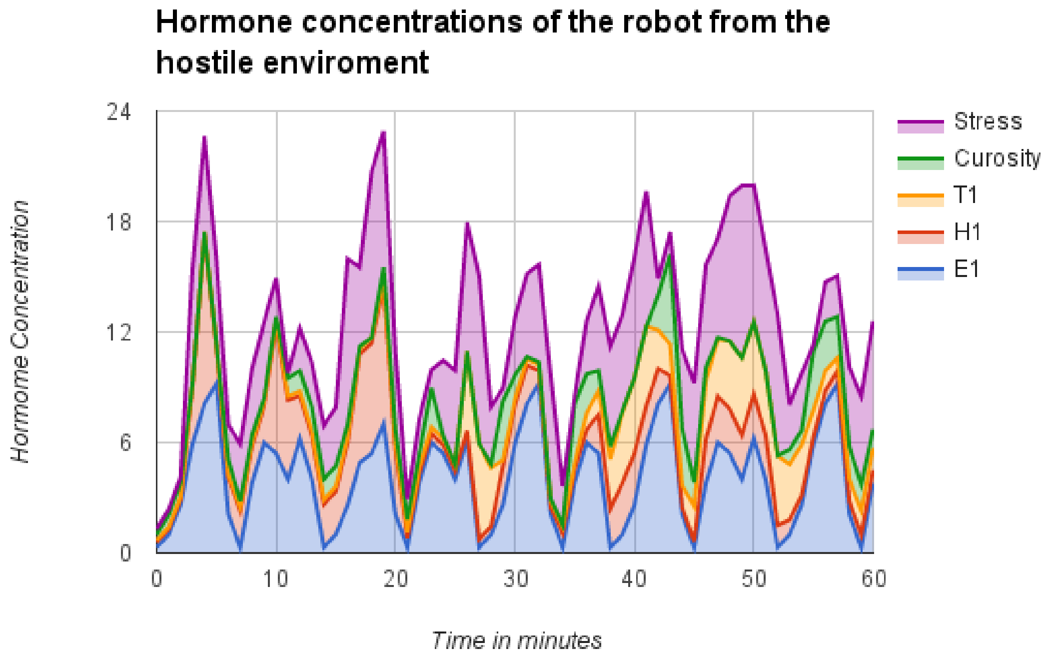


Fig. 6.4 The robot from the hostile environments hormone concentrations.

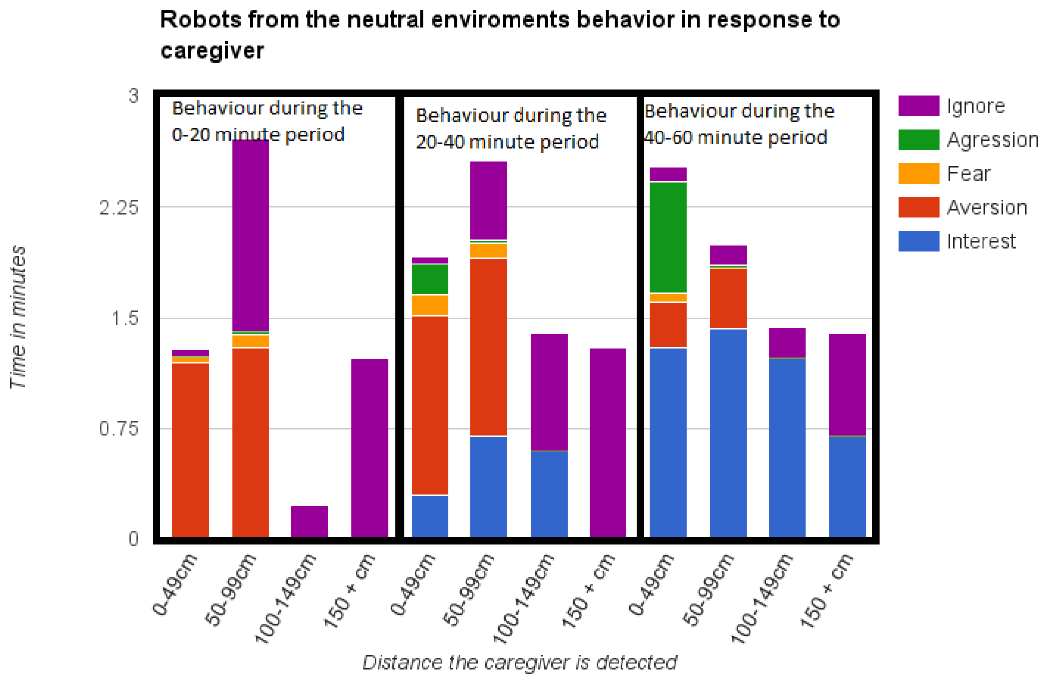


Fig. 6.5 The robot from the neutral environments behavior upon detection of the caregiver. The left box shows the response during the 0-20 min period, the center 20-40 min and the right 40-60 min

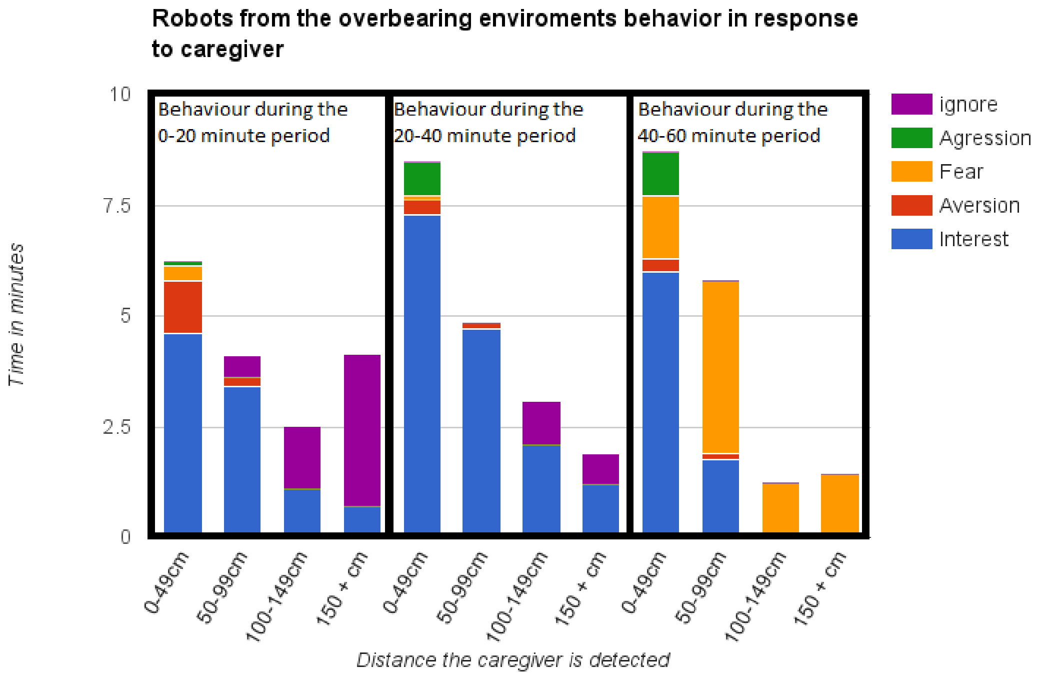


Fig. 6.6 The robot from the overbearing environments behavior upon detection of the caregiver. The left box shows the response during the 0-20 min period, the center 20-40 min and the right 40-60 min

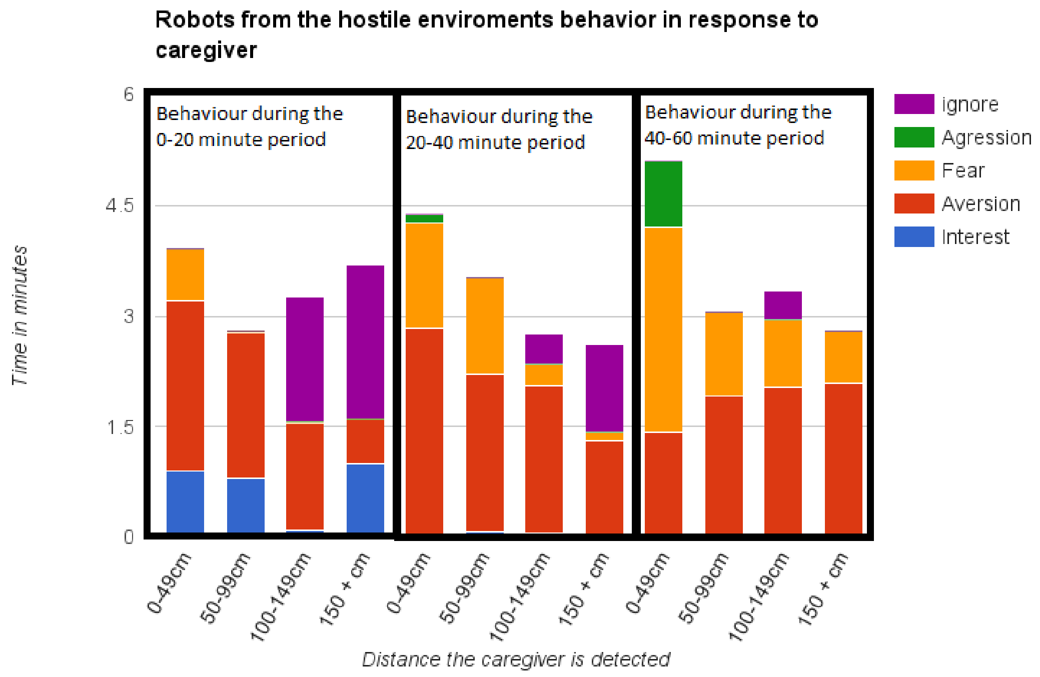


Fig. 6.7 The robot from the hostile environments behavior upon detection of the caregiver. The left box shows the response during the 0-20 min period, the center 20-40 min and the right 40-60 min

Almost all occurrences of the interest behavior occurred during the first 10 minutes of the run, before the robot had not adapted/learned about the hostile nature of this caregiver find the resources needed to maintain its homeostatic needs. During this period the robot mainly ignored the caregiver as shown in figure 6.5, the exception being if the caregiver moved within close proximity to the robot, in which case the robot would attempt to move away. The random walk and ignoring/avoidance of the caregiver took place due to the robot prioritizing the maintenances and recover of the survival related homeostatic variables.

The remaining 45 minutes: Exploration and interactions

By about the 15th minute the robot's behavior resembled less of a random walk and became more purposeful as the robot learned the location of the different resources and was able to move directly to them when needed. By learning the location of the resources the robot needed to dedicate less time maintaining it's homeostatic variables as shown in figure 6.2. The robot's ability to quickly sate its internal needs, and due to the lack of any other critical internal or environmental challenge resulted in the robot having a higher "wellbeing" during this period. This increased wellbeing can encapsulated by an increase in the curiosity hormone and reduced levels of both the stress and *eH* hormones as shown in 6.2. This concoction of hormones led to the robot having motivations to interact and explore with different aspects of surrounding environments with a particular interest in novelty.

The actual size of novelty the robot was attracted to was in direct relation to the current curiosity/stress hormone ratio. A high curiosity concentration and low stress concentration would lead the robot to search for object with greater perceived novelty potentially ignoring objects with a low novelty value, and vice versa with if the concentration were reversed. Once a suitable object had been found the robot would attempt to interact via gentle pushes, driving around or approaching it from various angle utilizing different sensors. These interaction lasted until either the robot needed to move to sate an internal variable, or the object no longer provided an appropriate level of novelty for the robot due to changes in the current hormone concentration and/or due to learning about the object so as to reduce its perceived novelty value.

In this experiment, in a majority of cases the reason for abandoning the interaction lied with a need to sate a homeostatic need or due to the decay of the curiosity hormone. However as the robot had started to learn about the object before breaking the interaction, the object became less novel in comparison to other aspects of the environment, reducing the chance of interacting with the same object multiple times. The meant that while the robot tended to interact with a large range of objects and aspects of the environment it never spent too long learning about each one, the implication of this is looked at in more detail in section 6.4.

Finally the interactions between the caregiver and robot, although the environment was set up as an independent scenario as previously described, interaction between both did occur. When these interactions did occur, excluding the first 15 minutes where the robot would always withdraw, the outcome was highly dependent upon the internal state of the robot. As the caregiver was the only other dynamic entente in the environment, the robot identified them as the largest source of novelty in the environment. Therefore the robot's interactions where driven by the current Curiosity/Stress ratio, a high curiosity ratio was associated with attraction and therefore interaction with the caregiver where as low ratio or high stress concentration was associated with the aversion or fearful behaviors.

When the interaction occurred as a result of the caregiver moving too close, the response of the robot would once again depend upon its current internal state. With a medium to high curiosity ratio the robot would tolerate both their presence and with a high ratio even contact[2]. However, if the robot perceived the caregiver's interaction as over stimulating i.e contact lasting to long (the length of contact tolerated was dependent on the curiosity/stress ratio and on average found to equal to a duration of $\frac{Curiosity}{Stress} \times 0.93$ seconds with a standard deviation of 0.07). the robot would attempt to withdraw.

However, if the contact occurred when the robot was "feeding" the robot displayed what could be considered an aggressive behavior by turning towards the caregiver and trying to bump them. A similar behavior could also emerge if the caregiver got in between the robot and a desired resource, leading to the robot deliberately bumping/pushing past the caregiver to reach the resource.

This aggressive both in this experiment and the other is correlated with a relatively rare occurrence of a high concentration of both the stress and curiosity hormones in a similar ratio (see figure 6.8, which in this environment would only occur during interruptions when trying to "feed". In the other two environments similar hormone ratios may also occur if the robot is exposed to a sudden stressful stimuli, while previously having a relatively high wellbeing . Due to the conditions needed for the aggressive behavior to occur, episodes of aggressive tend to be relatively, sudden and short lived with most occurrence of aggressive behaviors lasting on average around 2 seconds. The basis for the emergence of these behavior is the robot learning that by bumping the caregiver, they will move away as a natural response. This explains why the robot will deliberately bump the caregiver if they are between the robot and

---

[2] contact could occur if both the caregiver and robot where trying to reach the same part of the environment

a needed resource. When turning to bump the caregiver during feeding, looking at the wiring of the ENN the robot seemed to have associated the caregiver with a resource becoming unattainable, and therefore it is preemptively bumping the caregiver to make the move away. This wiring and learning is likely due to the caregiver blocking the robot's view of a resource when walking in front of it, and therefore robot associating them with the disappearance of a resource.

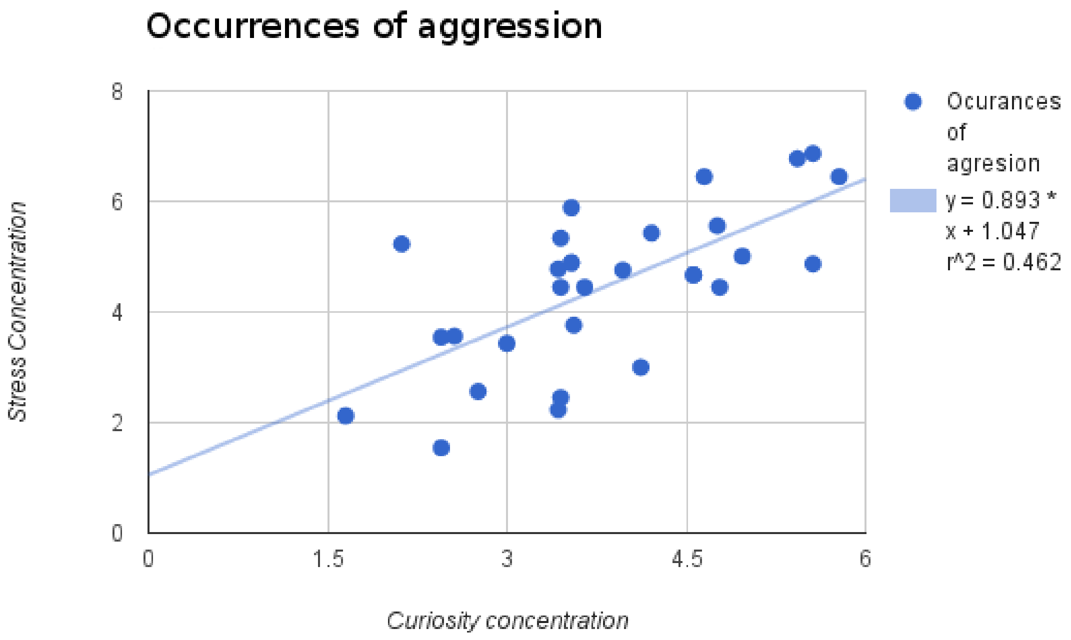


Fig. 6.8 Occurrences of aggression in relation to hormone concentration
As can be seen their is significant correlation between occurrences of “aggressive” behavior and the Stress/Curiosity hormone ratio

6.3.4 Overbearing caregiver

In the Second of the three experiments the robot was placed in the environment along with an overbearing caregiver. In this scenario the overbearing caregiver tailored their interaction (see table 6.1) with the robot to suit its internal state and to sate any needs of the robot. For example if the robot has a desire for energy the caregiver will bring an energy resource to the robot. A critical element to emphasis here is that all interaction between the caregiver and the robot are driven by the robot's motivations. The caregiver will not attempt to encourage any new desires or behaviors from the robot regardless of any perceived potential long or short term benefits. For example in the unlikely scenario that the robot's was overheating but showed no desire to reduce its internal temperature then the caregiver would not interfere or attempt to cool the robot. As with the robot in the neutral environment, this robot went through two distinctive phases with contrasting behaviors and motivations. The first corresponded roughly with two 30 minutes while second occurred during the remaining 30 minutes. The changes in the robot's state and behaviors can be seen in figures 6.3 and 6.6.

First 30 minutes: early interactions with the caregiver

Firstly due to the previously discussed early development conditions (see section 6.3.2) at the start of these experiments the robot had naturally developed a low tolerance to fast and sudden movements. Due to this intolerance, it was essential that if the caregiver wanted to interact with the robot without overstimulating it they needed to maintain slow and smooth movements . Failure to do so would lead to the aversion or fearful behavior. However, as the caregiver always maintained the same movements, the robot was never exposed.

In the first twenty minutes of this experiment due to the caregiver availability to provide resources and satisfying other needs as they emerged, the robot remained relatively sedentary. As the caregiver was quick to respond to any homeostatic deficits, the robot's motivation for movement was mainly due to rare occasions of novelty seeking behavior. Novelty seeking behaviors only emerged occasionally as the feeding interactions between the caregiver and the robot generated enough novelty to satisfy the robot's needs.

Finally in the early stages of the experiment interaction between the robot and caregiver was relatively one directional, always initiated and controlled by the caregiver and mostly consisted of the “feeding” interaction. However, shortly after the 18th minute the robot increasing began to attempt to follow the caregiver, actively trying to stay within a range of 10 to 30 cm. This can be seen in figure 6.6 where the robot is increasing showing the "interest" behavior and a reduction in the "ignore" behavior when the caregiver is detected

Second 30 minutes: later development and over reliance

By around the 28th minute the robot would now follow the caregiver whenever they were detected. This following behavior that emerged in this later period had similarity to the imprinting phenomenon seen in different animal species particularly nidifugous birds (see [82]). The trigger for the emergence of this “imprinting” is due to the positive feedback from interaction with the caregiver such as feeding.

With the emergence of this new behavior the interactions between the robot and caregiver increased substantially. Due to the robot's following behavior a majority of interaction from this point onwards was now initiated by the robot. This primarily consisted of the robot moving towards the caregiver and waiting for them to respond to one of it's needs, i.e to be "fed". Interestingly at this point, likely as a result of the extended interaction prior the robot had become entirely reliant on the caregiver. While in the first period the robot would attempt to find a needed resource before the caregiver intervened, now when the robot homeostasis levels dropped it would immediately move closer to the caregiver. This behavior would extended to the robot forgoing an easily accessible resource which allowed it to feed itself and instead move to the caregiver to fed by them. Their are two reasons fore this behavior

- Firstly uncertainty , the robot had limited experiences self-"feeding" and therefore the outcome of behaviors associated with moving directly to a resource were more uncertain in terms of recovery of a variable (signified by a lower neural plasticity along the pathway) than approaching the caregiver.

- Secondly affordance, as the caregiver could sate all needs, their overfall perceived affordance was further increased in comparison to a single resource which would only sate one need.

In a few cases when the caregiver was not visible to the robot but a resource was, the robot, would still rather search for the caregiver then replenish at the resource. It should be noted though that due to the caregiver responding promptly to the robot's motivations the robot would never need to search for long, and at around 10 seconds.

Another interesting phenomena that occurred during this phase was the robot's associating feeding with novelty. This occurred as feeding, the main interaction between the two was a novel stimuli to the robot due to caregiver's movement. Additionally the novelty of "feeding" was enhanced for the robot due to the lack of exposure to other novelty source. The implication of this is the ENN wiring in such a way that curiosity hormone would not only increase the motivation for novelty, but also increase the motivation to replenish its internal. This wiring meant that this robot could develop a motivation to feed without the presences of the *eH* hormones or a homeostatic deficit, with the robot using feeding as a proxy to get interaction and sate its desire for interaction. In this particular study, as this robot faced no punishment for overconsumption (internal variables could not go above their maximum), this was not a significant issue, however in other studies where this was not the case, this behavior would be highly detrimental to the robot's health.

A final interesting phenomena present in this environment occurred due to the robot becoming the primary initiator of interactions. In the early stages when the caregiver initiated interactions, they where able to perform the appropriate interaction as they had already checked the robot's motivations. However, when the robot imitated interactions the caregiver was not always immediately able to respond in the correct manner, as they first needed to see what the robot needed, leading to both a delay in their response and often needing to initially move away to the monitor. As the robot, due to prior experiences has learned to expect the caregiver to immediately sate its needs, and also has not been exposed to behavioral variation, the caregiver not acting as anticipated led to a secretion of the stress hormones a result of its low tolerance to behavioral variation from the caregiver as shown in figure 6.3.

The high concentration of stress lead to emergence of the fearful behavior (see figure 6.6) This behavior would usually encourages the robot to move away from the perceived stressor and to an area or object or perceived comfort. However, here due to the robot's reliance on the caregiver, the caregiver is the only known source of comfort. This meant that even though the caregiver was the source of the stress, the robot would still move towards them. Further the greater the variation in the caregiver's behavior the greater secretion of the stress hormone and the more intense the desire to seek comfort. In this scenario an increase in intensity of desire to find comfort leads to the robot trying to reduce the distance between itself and the caregiver, at relatively high level of stress secretion this will extend to the robot coming into physical contact. This resulted in the robot losing significant health (see figure 6.3) and as a result of this the caregiver spent a large portion of the final 10 minutes having to "feed" the robot the health resource in order to maintain the robot's health.

6.3.5 Hostile caregiver

In the third and final experiment the robot was placed into the environment with what we have called an hostile caregiver. Here the hostile caregiver would essentially act in an aggressive or unsuitable manner towards robot, with the behaviors of the caregiver including, fast stroking, sudden movement, chasing and unwanted contact (see table 6.2) . Over the course of the 1 hour experiment the robot perceived that it was exposed to a total of around 20 minutes of hostile interaction.

Early phases: First exposure to hostile caregiver

During the early phase of this experiment the robot's behavior when alone was similar to the behavior seen in the previous experiment 6.3.4 with the neutral caregiver. However, due to the unsuitable behavior of the caregiver, when they imitated a hostile interaction the robot's behavior quickly changed to avoidance or fear due to rising stress hormone concentration. Initially when the avoidance or fearful behaviors occurred the robot would move away from the caregiver in a manner descried in section 6.3.4. However, after around 10 minutes when these behaviors were trigged the robot would instead move towards walls or under desks. The reason for this change in behavior was robot learning that the walls and spaces under the desks provided protection against the caregiver interactions. Walls naturally protected one side on the robot will also make it harder (due to the caregiver's bio-mechanics) for the caregiver to imitated interactions. While under the desk the robot was almost completely protected from any hostile interactions.

Second phases: Preemptive behavior

The long term implication of exposure to unsuitable behavior started to emerge after the first 20 minute. At this stage the robot began to consistently engage in avoidance behavior preemptively as soon as the caregiver was detected rather than reactively after interactions where initiated (see figure 6.7). Additionally the avoidance behavior again began to change with the robot no longer simply moving towards a wall but now actively engaging in a wall following behavior, moving along the perimeter of the environment. This wall following behavior was observable during any period of medium to high concentrations of the stress hormone. With greater concentrations of the stress hormone the robot would move to and stop in areas which it perceived to be safe, either corners (Due to the detection of two walls) or under a desk. The robot would stay in these perceived areas of safety until either the stress hormone concentration had decayed enough or one of its homeostatic needs (normally energy) dropped low enough to force it to move to replenish it.

In cases when avoidance of the caregiver was not possible for example being tapped, the robot's stress concentration would rise to such a level that its motivation to "escape" would be large enough to suppress the motivation to avoid the caregiver and see it willing push past them . This pushing behavior was initially simply a “byproduct” of the robot trying to move away and not a true purposeful aggressive behavior as described in section 6.3.1 i.e their is no intent to hit the caregiver, simply to move away. However, as the robot found the caregiver would move away if bumped (for obvious reasons we did not want to engage in a physical contest with the robot), it learned that it was possible to end the hostile interaction with caregiver by making contact with them. To activate, the neural pathway associated with purposively bumping the caregiver needed the robot to have both a high concentration of both curiosity and Stress (see figure 6.8 and a low concentration H1 hormones, a relatively rare occurrence in this environment making purposive aggression rare.

Third phases: Exploiting the environment

After continuous repeated exposure to the hostile interaction the robot's behavior again began to change. Firstly the wall following behavior and hiding under desks became almost permanent with the robot only leaving the perceived safety/comfort of these areas in order to reach a homeostatic resource. The cause of this behavior is due to the significant up-regulation of the stress hormone receptor resulting in even basal concentrations of the stress hormone resulting in the avoidance behavior.

However, due the presence of the homeostatic variables it was impossible for the robot to stay under the desk or in a wall following behavior at all times. When the robot did need to leave these perceived safety, due to fear of the it open surroundings it would move at a considerable faster pace to minimize exposure. The implication here is robot risking overheating, to counteract this when the robot was in areas of perceived safety it would move at a slower pace keeping its temperature low and facilitating rapid burst of speed when needed. This translated as the robot moving at an average pace of $0.07m/s$ when in perceived safety and 0.3m/s when exposed. The fast movement also had another implication for the development of the robot. Due to the faster movements and therefore higher internal temperature in the open environment, if the caregiver attempted to try to interact with it in the open, the addition movement needed could result in the robot needing to stop due to a critical risk of overheating.

When this occurred as we were not expecting the robot to stop, interaction was disengaged under the assumption the robot had a technical failure. However, shortly after the caregiver disengaging and the robot recovering from overheating it would move away. Without us realizing what was occurring, this scenario repeated a few times during the 30-40 minute period of the experiment which allowed the robot to learn that by stopping it could "fool" the caregiver into leaving it alone. This can be seen in the 40-50 minute period where the robot abandoned the past avoidance behavior and would simply stop moving when the caregiver was detected. In this period stopping was no longer attributed to overheating and instead is a reactive behavior to the detection of the caregiver. By this point we had realized that

Table 6.3 Homeostatic variation of the three robots

Environment	Energy			Health			Temprature		
	Avg/SD	Tolerance	Desperation	Avg/SD	Tolerance	Desperation	Avg/SD	Tolerance	Desperation
Independent	68.2	45.6	37.4	81.2	64.1	21.4	11.2	17.5	32.1
Overbearing	91.3	94.6	94.1	97.5	98.1	95.7	4	4	10
Hostile	40.2	31	21	43.7	91	74.4	15.6	9.2	86.1

this stopping was not a technical failure and instead was a purposeful behavior by the robot. Once we identified the stopping as a purposeful behavior, the caregiver would continue to interact with the robot even once it had stopped. The first time the robot persisted with the play behavior for between 15-20 seconds before returning to it previous withdrawal tactics of wall following and hiding under desks. For each subsequent interaction the time playing dead decreased as the robot adapted to this tactic no longer being successful.

6.4 Comparing the robot's development

In this next section will look at how these three varying environments have affected the development of the robot. Firstly we will look the effect of caregiver's on the robot's homeostatic variables and motivations.

6.4.1 Effect of the caregiver on the homeostatic variables

The maintenance of the different homeostatic lies at the core of our model with development,behaviors and perception all influenced either directly or indirectly. Given the importance of maintaining homeostasis, we will firstly analysis and compare the varying effects that the different environments have had on this critical component which is shown in table 6.3.

One of the most immediate things shown by data is the variation in the average overall homeostatic levels. This has occurred primarily due to the robot developing different tolerances (these tolerance are unplanned and emergent) to homeostatic deficits in relation to their environmental conditions. Here tolerance, as shown in table 6.3, refers to the average[3] and subsequently behavior. For example when a homeostatic variable has decreased below its tolerance then we would expect the robot to activity begin to attempt to recover it. Desperation is special kind of tolerance which refers to the point in which on average the homeostatic

---

[3] Due to the way motivations are calculated the level the homeostatic variable needs to be reduced to before it begins to have an effects of the robot's motivation (See equation 2.5)

variable has dropped so low that replenishment becomes an overriding motivation effectively suppressing all other drives and motivations.

Tolerance and Desperation emerge as a function of the modulation of motivations due to environmental exposures as product of both the epigenetic mechanism affecting the internal drives (see equation 2.5 and equation 2.10), the learning process affecting the perceived affordance of different behaviors and or objects (ref section) and the current environmental conditions. Essentially the harder it is to recover and/or maintain a variable (i.e due to resource scarcity or from being hit) the lower the tolerance to any deficit will be.

Tolerance and desperation levels do not however appear to be correlated save that as expected desperation is always occurs at higher deficits. This asymmetry is influenced by relationship between, and the robot's ability to maintain the other two variables. For example in the hostile environment, table 6.3 shows that the robot will try to reduce is temperature when it reaches 9.2 yet it is only when its temperature reaches 91.4 that it reaches desperation, a significant gap in comparison to other variables. The reason for this is that the robot has adapted to needing to move rapidly move away from the hostile caregiver to avoid their interactions which this combination of tolerance and desperation is ideal for because:

- The low tolerance to temperature deficits means that the robot will try to keep it temperature low when their are no other pressing needs. By maintaining a low temperature, the robot can then engage in rapid movement when needed for longer before overheating

- A high desperation level will mean that the robot will tolerant higher temperature when other pressing needs are present i.e avoiding the caregiver. This means that the most cases the robot will not generate motivations to reduce its temperature (slowing or even stopping) until the caregiver is no longer a threat (i.e., the robot reaches a safe position such as under the desk)

Therefore the differences between tolerance and desperation essentially allows for the emergence on unique traits in regards to the environmental conditions in which the robot has developed in.

6.4.2 Change in homeostatic variables of the robot from the neutral environment

Without any significant environmental challenge the robot from the neutral environment developed relatively high tolerance to the homeostatic deficits. Out of the three homeostatic variables the robot was least tolerant to energy deficits, unsurprising since this is the variable the robot has least control over, (I.e the battery continuously decrease regardless of the robot's behavior).

6.4.3 Change in homeostatic variables of the robot from the overbearing environment

The robot developed with the overbearing caregiver as would be expected maintained very high high homeostatic levels through this experiment due to the feeding interaction. Less expected however was the robot's extremely limited tolerance to said deficits. The reason for this turns out to be two fold.

- Firstly as this robot has been developed in an environment where high homeostasis is the norm, it has in essences developed no concept of "hunger/damage" and has adapted to high homeostasis as being the norm.

- Secondly as mentioned in section 6.3.5 the interactions from being feed a resources is the main source of novelty for the robot. As discussed in 6.3.4 it appears that the robot's motivation to consume is not purely driven by a homeostatic deficit but also due to a desire for interaction and novelty, leading to increased motivations to feed. Further backing up this assumption is also shown in table 6.3 where the tolerance to temperature, the homeostatic variable that is not recovered via interaction with the caregiver is larger than the other two variables (although still much smaller in comparison to the other environments). In other words since this variable is not recovered via interactions, the ENN has not wired in such a way that the curiosity hormone effects the motivation.

  Fig temp and curiosity

6.4.4 Change in homeostatic variables of the robot from the hostile environment

The robot from the hostile environment had three main interesting homeostatic occurrences.

- Firstly, it has a much lower tolerance to health loss indicating that taking damage was its biggest concern. This would be a valid observation from the robot since the hostile caregiver main interactions would result primarily in health loss.
- Secondly, the robot displayed multiple short term temporal adaptation through significant over consumption of one type of resource. For example the robot would over consumer on energy allowing it hide for longer period in perceived safety areas, i.e similar to hibernating. Alternatively in period when the hostile caregiver was more active the robot would over-consume on health to allow it withstand more hostility from the caregiver (ref section remember to add section of homeostatic variables interaction)

  figure for this
- Thirdly, the robot developed an unusual big gap between its temperature tolerance and desperation levels. The combination of these two factors meant that the robot meant that the robot would try to maintain a low temperature, but would suppress this desire in the presence of other motivations. Essentially this behavior kept the robot in state of preparation to allow it rapidly move away in the presence of negative stimuli (i.e the caregiver)

6.4.5 Effects of the environment of the robot's learning

The next comparison we make deals with the learning experience of the robot seen in 6.4. As previously discussed in chapter 5 learning occurs as the robot interacts with different objects and environments as part of the curiosity driven motivations. From these interactions the robot is able to learn the affordance of the said object or environment by learning the effect that they have on its internal variables e.g if the robot from the hostile environment stays in the middle of the room it is more likely to be exposed to hostile interactions fro the caregiver.

Table 6.4 The learning of the three robots

	No of uniq obj/ KNW	No OBJ CMP	CMP ACC	AVG Neural strength	Neural activity	Response		
						C	F	N
Over	56/ 97% | SD 1.1	32	42%	100%	15%	2	94	4
Independent	4023 / 41% | SD 8.2	21	89%	41%	43%	58	12	30
Aggressive	1497 / 62% | SD 22.9	212	71%	83%	100%	6	80	14

6.4.6 Learning experience of the robot from the neutral environment

Due to the lack of significant danger and challenge in the environment, the architecture has lead to the emergence of an outgoing phenotype, constantly exploring and searching for new and interesting phenomena within its environment. As we mentioned in sections 6.3.3 the independent robot interacted with a large range of different aspects/objects of its environment, but it did not spend too long interacting with any particular one. This is reflected in the robot's learning, where we can see that it has identified a high number of different objects or aspects of the environment, yet only "believes" it has full knowledge of 21 of them.

The benefits of this behavior is potentially that by interacting with more objects the robot is likely to discover new ways in order to adapt to and strive in its environment. One such example is the robot learning that resources could be found easily through the use of landmarks. Unfortunately in this environment their was only limited opportunities to test this hypothesis, however in chapters 6 we do look into more details about the potential for this architecture to learn and exploit unique characteristics of its environment.

A final interesting point to make here is the way the robot reacted when an object behaved unexpectedly. While the robot from the Hostile and overbearing environment overwhelming respond fearfully, withdrawing if possible the robot from the independent environment treating variation in object behavior as a novelty and it perceived attractiveness increased.

6.4.7 Learning experience of the robot from the hostile environment

Because of the hostile nature of this environment lead to constant high stress concentrations this robot rarely looked for new sources of novelty due to aversion of novelty. The implication of this as can be seen in table 6.4 is that the robot has identified a much smaller number of objects. However, of those identified the robot has a deeper learning in comparison to the other robots. Specifically this robot learning revolved around critical aspects of its environment. This is due to the stress concentration not only reducing the perceived attractiveness of unknown objects due uncertainty, reducing the chance of interactions, but also affecting the robot's focus or attention. Essentially an unplanned ability, it appears that when the ENN and Neuromodulation systems are combined experiences associated with extreme stress or curiosity are learned about faster due to the increased intensity of neural signals sent along that pathway. There is a trade off however to this rapid learning with the quality of the learning, with the robot more likely to learn of inaccurate affordances or characteristics

6.4.8 Learning experience of the robot from the overbearing environment

Lastly the robot with overbearing caregiver, as would be expected due spending it time almost exclusively interacting with the caregiver learned about the smallest number of objects in its environment. However as the robot was constantly exposed to the same objects it did believe that it had learned fully about half the objects it had been exposed to. Interestingly though, half of the objects the robot thought it had fully learned about actually behaved in a different way than it expected. This is likely due to the way in which the caregiver introduced object to the robot. For example due to the caregiver bringing the energy resource to the robot, the robot had learned that these resources moved towards it, rather than them being static objects which the caregiver was manipulating. Additionally as previously mentioned (see section ref) this robot appeared to develop more specific or tighter criteria for the expected behavior or affordance of an object . Meaning that even a slight variation might be enough to make the robot think the object behaving in a new manner.

6.5 Measuring the suitability of development

While we have demonstrated that robots exposed to different environments have developed differently we have not yet considers if these variations are beneficially to the robot. In order to so, each robot after it had finished the one hour experiment in its own run, was placed in the other two environments for 10 runs of 10 minutes in each environment. With the outcome of this seen below (at the end of each 10 minute run the robot was reset to the same state as when it finished the one hour in its own environment).

6.5.1 Neutral caregiver

The robot from the overbearing environment

As would be expected the robot from the overbearing environment which had developed to rely on the caregiver for feeding was unable to operate when placed into the environment with the neutral caregiver. Here this robot "died" early due to lack of energy on every run. This robot predominately searched for or tried to interact with the uninterested and unattentive caregiver.

The robot from the hostile environment

The Robot from the hostile environment managed to survive in the neutral environment with limited problems. While this robot did have lower homeostatic levels than the robot from this environment they never dropped to what we would consider critical levels. The main issues as would be expected was simply the robot engaging in unnecessary avoidance or fearful behavior when the caregiver posed no threat, resulting in inefficient movement.

6.5.2 Overbearing caregiver

The robot from the neutral environment

The robot from the neutral environment initially performed well when placed with this caregiver. At the start the caregiver's "feeding" resulted in the robot maintaining high homeostatic levels while also being able to move about and explore the environment. However after around 6 minutes on average due to the faster learning of this robot (see section 6.4), the caregiver was soon deemed to have a low novelty values and therefore became unattractive to the robot when it had a high concentration of the curiosity hormone. As the robot's curiosity concentration was almost always high due to the caregiver interactions resulting in a high wellbeing, at this stage the robot entered a behavioral loop. Whereby the robot would constantly try to get away from the caregiver when its wellbeing was high (the caregiver of course following resulting in a chase), before turning back to them to be "feed" when homeostatic variables dropped.

The robot from the overbearing environment

The robot from the hostile environment performed exceptionally badly when placed with the overbearing caregiver. Due to its past negative experiences with the hostile caregiver, this robot would become stressed when a caregiver was detected. Here the constant interactions of the overbearing caregiver who may of had the best intention of the robot in hand, resulted in exceptional high concentrations of the stress hormone resulting in the robot spending all its time trying to move away from the caregiver and ignoring its own survival needs resulting it death from lack on energy, health (due to collisions) or overheating (trying to move away) [4].

While as we discussed in section 6.3.4 the robot developed in this environment did begin to show many negative traits, i.e poor learning it still performed better in this environment that the other two robots.

6.5.3 Hostile caregiver

The robot from the overbearing environment

Much like in section 6.5.1, the robot from the overbearing environment would simply "die" without the overbearing caregiver. Due to hostile and unsuitable behavior of the caregiver, the robot displayed the same behavior as in section 6.3.4. Whereby, even though the caregiver was causing the robot to become stressed, as they were the only source of comfort known, the robot would try to get closer and coming into contacts with the caregiver. This behavior resulted in either the robot dying early due to a lack on energy or due to a lack of health from collisions

The robot from the neutral environment

The neutral robot showed some ability to survive in this environment. When the caregiver tried to interact in a hostile way this robot would either show an avoidance behavior if it had a high ration of stress to curiosity or the aggressive behavior with high ratio of curiosity to stress. In most this fight or flight response proved suitable to robot's current environment. However, on three occasions the robot did die. On the first occasion the robot choose to

---

[4] While the caregiver did try to use calming interaction to reduce the stress, due to this robot's perception of the caregiver the only option would of been to simply leave the robot alone, which was not an allowed option under these experimental conditions

fight with low health and the contact from bumping the caregiver resulted in its death. On the other two occasions as this robot did not develop the same stress coping mechanism of the robot from hostile environment (hiding under desks or wall following, see section 6.3.5, which allowed the stress hormone concentration to decay between caregiver interactions) this robot's stress hormone concentration kept building up resulting it such high concentration near the end of the run that it becomes completely risk adverse, remaining stationary at the side of the environment for increasing periods and forgoing the recovery of its variables needs. If the experiments is carried for an additional 10 minutes then all neutral robots dies from one of the previous two causes.

6.6 Conclusion

In this chapter we have tested the ability of Neuro-modulatroy system from chapter 4, and the Emergent Neural Network (ENN) from chapter 5, to work together to create an adaptive learning system. Both systems have the potential to work interdependently to allow the robot to adapt to its current environment. However as we discussed in their relevant chapters, both systems when used independently have weakness. The Neuro-modulation require the designer to pre-program the robot with the affordance of aspects and object present within the environment. As we demonstrated in chapter 2 affordances can be pre-programed for simple sterile environments with relative ease. However as environments become more complex the challenges of providing this information grows exponentially.

When combined both the ENN and Neuro-modulatory system seem to provide the desired adaptive learning experience for the robot. With each system able to compensate for the shortcomings of the other. For example the robot from the neutral environment was shown to be able to regulate its learning and constantly seek out novelty while still maintaining its survival related needs. However, the robot from the hostile environment which was exposed to a harsher environment, was more risk adverse reducing its interaction with novelty. Even though this robot was not able to spend as much time interacting with novelty it was still able to learn about critical elements. This occurred as a result of the mechanism leading the robot to be attracted and therefore focus on these key elements.

Additional experiments were also carried out to demonstrate that this variation in learning and behavior adopted by the different robots was beneficial.To test this each robot was placed into the others environments. Here we saw that in every case the robot that originated from that environment was better suited to those conditions than the other robots.

One unexpected occurrence in the experiments however was the emergence of unplanned developmental phases. Each robot had at least two phases with the robot's showing behavioral variation in each phases. We had all ready discovered in chapter 2, that the architecture allows for the emergence of unplanned behaviors. However these phases may suggest the combination of both ENN and Neuro-modulatory systems not only allows the robot to develop unplanned behaviors, but also build upon them. Leading to the emergence of more advanced behaviors as the robot develops.

The exception to this however is the robot from the overbearing environments where the robot's wellbeing decreased in the later phases. One explanation for this could be the lack of independence and exposure to different experiences essentially stunting the growth of the robots ENN and overall development. Crudely we might describe this as a use it or lose it like scenario, whereby due to the consistency and lack of variation in interactions the robots network has reduced plasticity inhibiting its future adaption (see section 5.2.3).

In the next chapter (Chapter 7 we will explore both these later hypothesis. The "ability for the architecture to build upon itself" and the "role of stimulation in the development of the robot".

Chapter 7

Novelty and cognitive development

In the previous chapter, we showed that the combination of the ENN and the epigenetic neuromodualtion mechanism led to the emergence of a learning system that modulated by the robots current and past histories. This meant that the robot would interact and therefore learn about its environments in an adaptive and appropriate manner. For example the robot which had developed in a dangerous environment was much less "outgoing" and would only interact with new or relatively unknown objects when all other needs were sated. Of-course if the risk or danger associated with the object was deemed too great then this robot under no circumstances would approach. As we demonstrated in it original "dangerous" environments this withdrawn robot significantly outperformed other robots in both qualitative and quantitative analysis. Similarly the other robots all outperformed each other when tested in their original environments.

However one of the interesting phenomenon that we began to see emerge during these prior experiments was the effect that early life sensory motor experiences had on each robot. Specifically as we noted in section 6.6 a robot's behaviors and adaptation seemed to increase the greater its exposure to different forms of sensory motor experiences. Specifically the outgoing robot which was exposed to a greater range of sensory-motor stimuli, specifically during it early life seemed to develop a more responsive neural network as subsequently appeared to develop better learning capabilities through the experiment.

In this chapter we investigate what roles sensory motor-experience play in the development of our epigenetic emergent neural network.

7.1 Looking into the unplanned development in the robot

To look into the previously described emergent developmental stages we designed a range of different experiments to try to understand this process.

Firstly in order to looking into the effect that the environment had on these developmental stages we designed three different environments. Still using the Koala 2 robot three different robots using the same architecture where then placed into each environment.

These environments consisted of;

- A base/standard environment (see section 7.1.2).
- A novel environment (see see section 7.1.3).
- A sensory deprivation environment (see 7.1.4).

Each robot was allowed to develop in their environment for a period of 50 minutes. Before each robot was placed into their respective environment however they spent 10 minutes with a caregiver in a pre-training phases (see section 7.1.1) where they were exposed to important aspects of the environment in a controlled manner, such as the energy resource, allowing them to learn its affordacnes i.e., it can recover energy. Without this, unless a robot was lucky and found and learned about a resource early they would likely "die" early into the experiments. Therefore the purpose of this pre-training phases is to teach the robots the basics skill they need to survive. In addition, this give us an opportunity to look into the early behaviors possessed by the robots before they begin developing. Providing us with a better understanding of how the have developed.

In the three environments, the robot would have access to two sources of each type of resource; in the third experiment, this meant that the resources were placed inside the box along with the robot.

7.1.1 Pre-training

Before the robots were placed into their different environments each spent the first 10 minutes of their "life" with a caregiver. This caregiver provided an identical experience for each of the robots with the primary purpose to teach the robot the critical aspects needed to survive, such as helping the robot learn how to recover from homeostatic deficits by bringing a resource

Fig. 7.1 Different aspects of the environment used during the experiments.
1) shows a panoramic picture [1] of the standard open environment used in sections 7.1.2 and 7.1.3. 2) shows an example of one of the novel structures used during experiment 7.1.3. 3) shows the koala robot used during this work and the cardboard box the robot was placed in to create a sensory deprivation environment. 4) shows two AIBO robots used as novel objects in the test described in section 7.2.2.

to the robot and allowing it to explore it with its different sensors[2]. This period essentially consisted of the caregiver sating the robots needs by bringing the relevant homeostatic resource to them. During this period these robots behaviors' were essentially, entirely driven by exposure to stimuli. In some ways, this basic behavior is similar to the so-called reflex acts of a newborn. At this stage, both newborns and the robots display many "reflex" behaviors;

---

[2] A video example can be seen at http://www.emotion-modeling.info/videos, titled Sensorimotor Experience and Affective Development in an Epigenetic Robot

for instance a newborn will “grasp” objects placed into their hand or suck an object placed against their lips; our robot’s “reflexive” behavior will generally see it move towards or away from (attraction vs repulsion) different environmental stimuli.

In this first phase, the interactions between the robot and the caregiver resulted in the emergence of five main reflex behaviors.The first three of these behaviors occur due to the innate design of the homeostatic variables, these being Attraction and Repulsion, Avoidance, Recoil. Attraction and Repulsion emerged when the caregiver feed the robots by placing a relevant homeostatic variable in front of it. This “feeding” by the caregiver, saw the robot would move towards the caregiver when hungry and then away when sated. Avoidance emerged when the caregiver moved too close to the robot, resulting in the robot moving towards into an area of more space. Recoil emerged when physical contact occurred, unlike with the avoidance behavior, here the robot will prefer to move in an opposite direction to the stimuli rather than simply towards more space.

The final two reflex behaviors seen during this period are slightly different. The Exploration behavior emerges due to a combination of the first three behaviors. The attraction behavior gives the robot the motivation to move forward while the avoidance and recoil lead to a motivation to avoid collisions. Finally“localized attention” the last innate behavior seen during this period is based partly on learning and emerges around the 8-minute mark. This behavior sees the robot turn to face a moving object that is roughly within a 30cm range. The basis behind this behavior can be traced to the fact that the robot at this stage associates movement with the presence of the caregiver[3] and therefore the impending “feeding”, which can only occur if the robot is facing the resource (and hence the caregiver holding it). At the end of this initial period the caregiver would leave the environment and be outside the robot’s view.

7.1.2 First Experiment: The standard environment

In the first of the three experiments the robot was placed in our open lab environment shown in figure 7.1. For this experiment, the robot was given free roam of our lab with only limited changes to the environment made. These changes include: (a) the use of plywood borders to block access to “problem” areas where the robot’s sensors and actuators would be unsuitable, and (b) the placement of resources. Additionally, blackout curtains were used here to block natural light entering through the windows of the lab.

---

[3] Although the robot did posses color vision, at this stage, perhaps due to environmental noise or due to slower development of vision, the robot at this stage relied on movement to detect the caregiver.

First Experiment: minutes 10-20

During this period of the experiment the caregiver was removed from the environment, and with him the feeding interaction between the caregiver and the robot. From now on, in order to maintain its homeostatic balance, the robot would need to seek out the different resources scattered throughout the environment. Resources were placed in manner in which they could be clearly seen by the robot – four resources, one in each corner, alternating in type – with the aim of causing it to move around the environment in order to experience different sensorimotor stimuli. The prior 10-minute exposure to resources through the caregiver's feeding was enough for the robot to have begun to learn some of the key features of the different resources such as their shape, color and size, to allow the robot to detect them.

The immediate challenges that the removal of the caregiver presents to the robot are threefold. Firstly, the robot must be able to manage conflicting needs e.g, if it chooses to replenish energy it must at least temporarily forgo reducing its temperature or replenishing its health. Secondly, the robot needs to develop tolerance so its consumption pattern – particularly to what level it can let a homeostatic variable drop before replenishing – is appropriate for the current environment. Thirdly, the robot must adapt its sensorimotor behavior – how fast to move and when to turn to avoid collisions – to the current environmental conditions.

At the start of this period, the robot was highly sensitive to its internal needs –attempting to replenish any variable that was roughly below 90%. Due to spacing of the resources the robot was often able to see at least one of each type at any given time, and therefore at this point it did not search the environment when a deficit occurred but rather moved to the nearest perceived resource. This movement was often inefficient as in many cases a closer resources was located outside its immediate field of view, either to the side or behind. However, at this point in time the robot's behavior was still largely reflex-driven – seeing the resource led to the robot moving towards it. When two homeostatic variables were low and the required resources could both be seen, the robot's choice of which variable to recover first would be determined based upon the size of both the internal deficit and the detected stimuli. A problem with satisfying needs in this manner is due to a combination of noise – the perceived size the of external stimuli would fluctuate – and homeostatic variables not decreasing linearly or at an equal rate, the robot's intrinsic motivations would fluctuate, and hence its "goals" and executed behaviors often changed before a need was sated as shown in figure 7.3.

First Experiment: minutes 20-30

The inefficiency in the robot's behavior after the withdrawal of the caregiver initially lead to the robot having issues in maintaining homeostasis. However after the robot had been sufficiently exposed to its environment and the epigenetic mechanism began to regulate hormone receptors, its behaviors became more appropriate, leading to the robot on average being able to recover a homeostatic deficit 54% faster. This can been seen in figures 7.2 and 7.3 which show respectively the change in the robot's movement patterns and motivations.

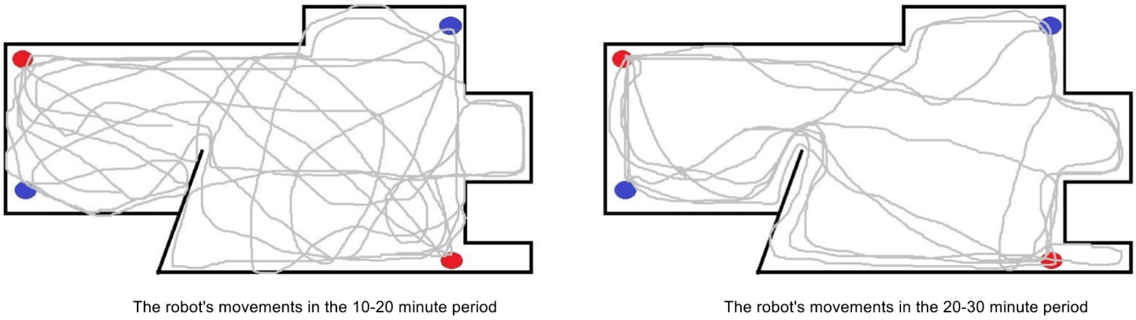


Fig. 7.2 A Comparison of the movement patterns of the robot in the 10-20 minute period (left) minute period of the experiment and in the 20-30 minute period (right). The red and blue dots represent the location of the health and energy resources respectively. Here we can see that in the later period the robot movements become more purposeful moving directly between the different resources. It should be noted that the movement maps were created using data from the robot's wheel speeds, rather than an overhead recording, therefore there may be some discrepancies.

As shown in figure 7.2 and 7.3, the robot's movements have become much more efficient for its environment, as it now moves more directly between the resources with limited motivation or behavior switching. This occurred firstly as a result of a change in tolerances to homeostatic deficits. As the robot was exposed to consistently lower but stable homeostatic variables due to needing to feed for itself, it soon became tolerant to these lower levels through the epigenetic mechanism. This resulted in reduced urgency in replenishing its internal variables, to the extent that they would now need to reach an average level of around 60% instead of the previous 90% before the robot would become motivated to replenish them. As a consequence of the reduced need to replenish the homeostatic variables of energy and health, the robot was no longer under such internal pressure to move quickly between the resources, allowing the robot to reduce its overall speed, resolving the issues of overheating and increased collisions associated with faster movement in the previous period. Additionally, while the robot maintained a relatively constant speed in previous periods, slowing down only to consume or due internal overheating, now the robot began to modulate its speed to

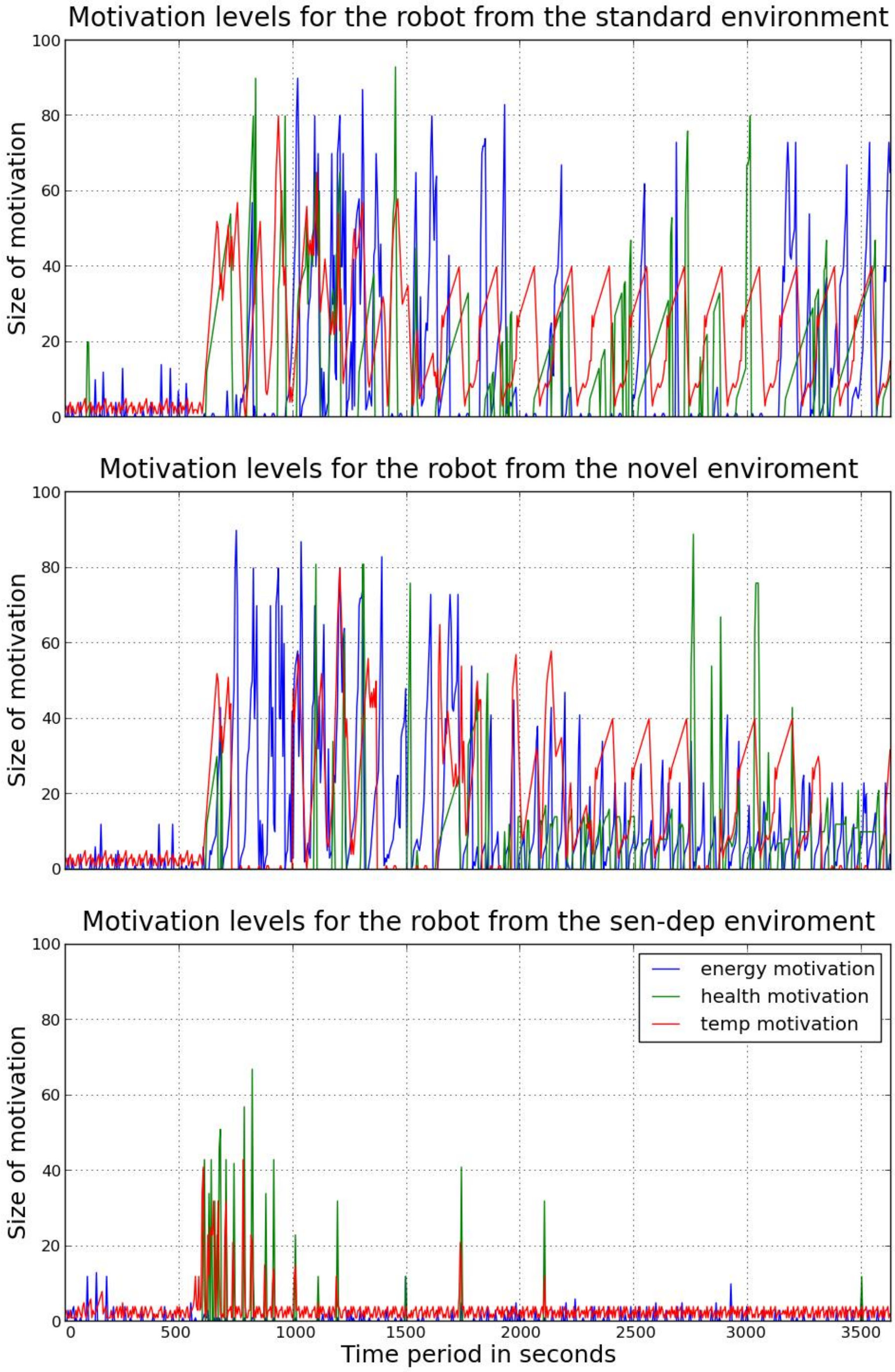


Fig. 7.3 A comparison of the three main motivations
A comparison of the three main motivations (replenishment of a homeostatic variable) for each robot during the experiments. As can be seen the early period changes and growth of motivations are much more volatile. This leads to increased occurrences of rapid behavior switching. Due to the volatility of change, this can lead to both inefficiencies and missed opportunities, i.e constantly moving between two resources without (fully) feeding.

match the environmental conditions. For example, the robot would move slower near the edges of the environment where it previously had collisions, and faster in the open middle areas.

This period represented an important time in the robot's development. As shown previously, during the early stages of this experiment when the robot was first exposed to this environment its behavior was almost entirely reflex-driven. However, due to motor-stimulation the robot's behavior has started to become adaptive, taking into account the current environmental conditions and its own physical body.

This period therefore potentially bears some similarities to the concept of primary circular reaction in infant development. Much like with infants at this stage, here the robot's focus is on the effects that its behaviors had on its own body – for instance, developing appropriate movement speeds and understanding and adapting to the restraints of the levels of its homeostatic variables. Similarly, for both the robot and the infant, behaviors categorized as primary circular reactions emerge as accidental discoveries [111, 122].

First Experiment: minutes 30-40

During the first 30 minutes or so the robot had begun to adapt its behaviors, focusing on maintaining homeostasis. This occurred by through the emerge of behaviors which have similarities to primary circular reactions. However at this point the robot began to show the emergence of more complex behaviors that could be considered similar to secondary or even tertiary circular reactions, as we will discuss in more detail below. At around the 33 minute mark, due to the robot's previously discussed reduced need for, and increased efficiency in, maintaining homeostasis, the robot spent a much smaller proportion of its time attending to homeostatic needs, showing a reduction from 93% of its time actively searching for resources in the first 30 minutes, down to 59% in this period shown in figure 7.4. This reduction in time needed to maintain homeostasis provided the robot with the opportunity to *explore* and interact with other aspects of the environment. During this period of exploration, using the previously discussed novelty mechanism (see section 5.2) the robot's motivations were determined by both the internal and external environment. Such exploration would take different forms, depending on hormonal levels. With high levels of the nH_c, which is associated with positive stimuli and a good level of homeostatic variables, the robot's attention was focused on the novel aspects of the environment. These novel aspects tended to be objects or areas that the robot had limited knowledge of, and/or objects that had some perceived uncertainty or danger as to the outcome of any interaction. In contrast, with

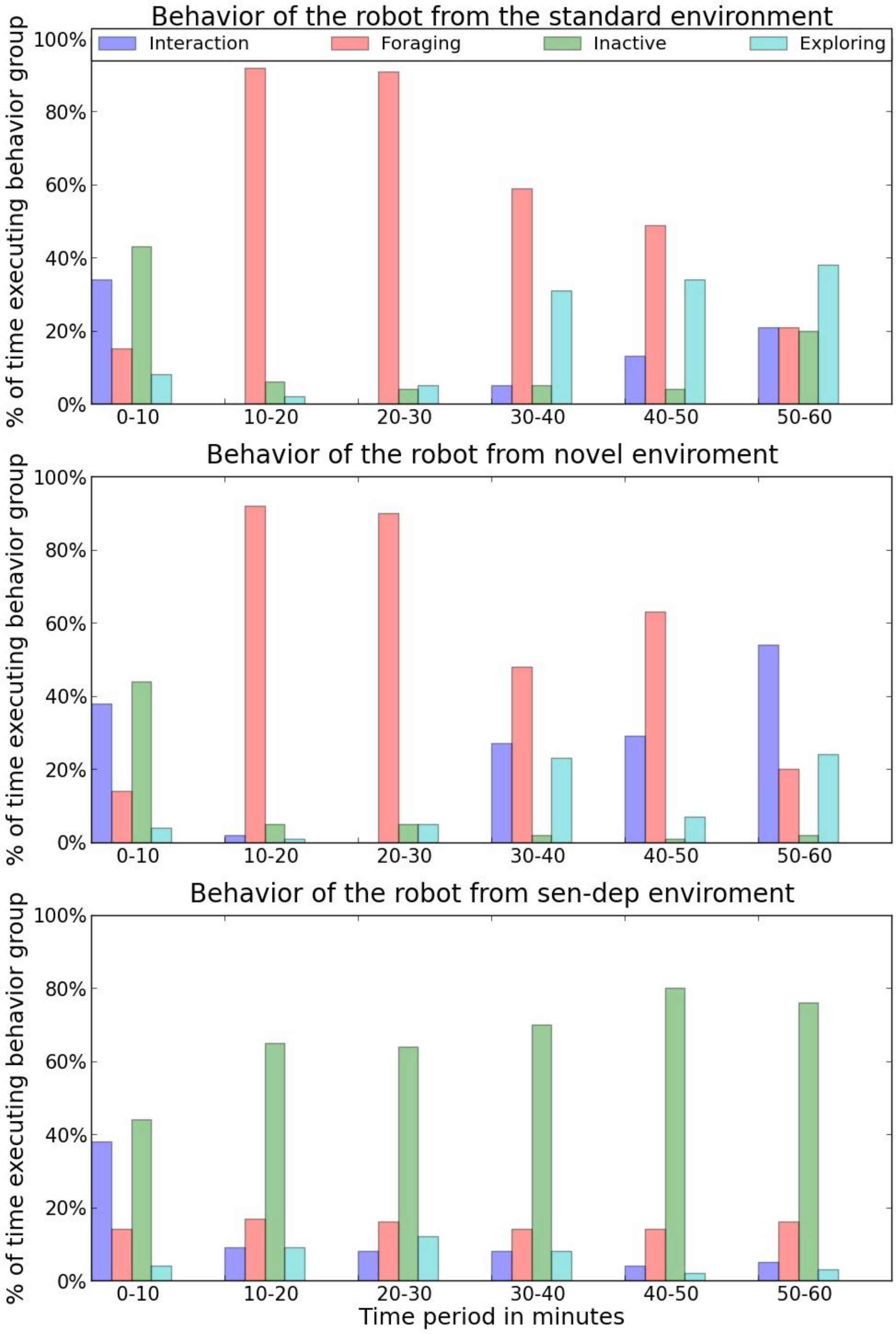


Fig. 7.4 The robots' behaviors

The type of behavior executed by each robot during each 10 minute period. As previously stated the robot has no explicit behaviors, instead behaviors executed by the robot have been classified into four general groups. Interaction includes any purposeful movement towards or contact with an aspect of the environment, foraging refers to any behavior that deals with the recovery of a homeostatic variable, this includes consuming, moving towards and searching for a resource, exploration includes any movement based behavior, while finally inactive is any period where the robot remains stationary without consuming or engaging in interactions higher levels of the nH_s, which is associated with negative stimuli, over-stimulation and low homeostasis, the robot is more attracted to, and will interact with less novel aspects, such as those it already had some understanding of or perceive to be safe, e.g., the walls of the environment due to their static nature. In cases where very high levels of the nH_s were present, the robot would simply move to an area of perceived safety and only leave when the nH_s levels had decreased sufficiently.

This period represented an important stage in development of the robot for two critical reasons. Firstly, during this period the increased exploration is strongly linked to the growth of the ENN (see section 7.2.1). Secondly, this exploration and interaction represent an opportunity for the robot to further understand both its own body and the ways in which it can influence its environment. Due to the relatively static nature of this first environment – most objects were either immovable or too large for the robot to meaningfully interact with them – interaction was relatively limited; it consisted for the most part in pushing an object for a few seconds, before learning that the only outcome of this behavior was a reduction in its health due to the contact, thus reducing future attempts to interact with the said object. However, around the 38th minute, the robot found the homeostatic resources which consisted of small plastic balls, light and easy to push, and therefore the robot was able to create an interesting novelty experience for itself by pushing[4] the balls.

First Experiment: minutes 40-60

During the latter stages of this experiment, due to improved efficiency in recovering homeostatic deficits, the robot spent most of the time either idle or interacting with homeostatic resources. Initially this interaction consisted of small pushes that took place over a period of around 10 minutes. The motivation for the robot to push the balls was twofold. Initially the pushing was curiosity driven, as the robot tried to learn what the pushing resulted in. After around 5 minutes, however, the pushing became novelty driven, caused by the new element of motion, as mentioned in the previous section. As expected in our model, due to the high novelty that resulted from pushing an object, the robot would only “purposefully” push objects when it had high ratio of nH_c to nH_s concentration.

This emergent behavior presents some similarities with ideas of secondary circular reactions. For example, a child using a rattle and our robot pushing the ball share the fact that the agent is beginning to notice and explore that their actions and behaviors can have interesting effects on their surroundings. Similarly, later the behavior where we see the robot pushing the ball in

---

[4] The novelty here stemmed from the movement of the ball rather than the actual action of pushing an object.

order to create a novelty source has similarities to progression of secondary circular reaction to coordinated secondary circular reaction, where the robot is now demonstrating the ability to manipulate an object to achieve a desired effect.

We observed another interesting phenomenon at around the 47 minute mark, as the robot seemed to develop a search strategy while looking for resources. Previously, when searching for a resource, the robot would randomly explore its environment; however, at this point, the robot began to show some strategy in its search, since instead of the random exploration, it would now move to the walls and follow them to search for the resources, which were placed near the corners of the environment. The emergence of this behavior further reduced the average time spent searching for a resource from the previous 59% down to 47%. As time went on, this behavior continued to develop and the robot began to learn to associate certain easily identifiable landmarks in the lab, such as a blue screen or a cupboard, with the presence of a particular resource. This ability greatly improved the time needed to find a resource, further reducing the average time spent searching for resources down to 21%. This behavior might suggest that the robot might have developed some notion of "object permanence" beyond learning a simple association between resources and landmarks. In order to investigate if this could be the case, we carried out the experiments reported in section 7.2.3.

7.1.3 Second Experiment: The novel environment

In the second experiment, we developed the robot in an environment very similar to the one used in the first experiment, with the difference of the inclusion of a range of different novelty sources. These included light movable objects arranged in the various shapes and patterns, as shown in figure 7.1, and two small Khepera robots that moved around randomly. If, at any point in time, any of these object were knocked over (e.g., due to the Koala robot's interactions) or stopped functioning as intended, the caregiver would replace or reset them as soon as the robot had moved away.

Second Experiment: minutes 0-30

As we would expect, in the early stages of this experiment the exposure to additional (compared to the first environment) sources of novelty had no real effect on the robot due to its preoccupation with maintaining homeostasis. Apart from the need to avoid the two additional randomly moving robots and the additional novel objects, the behavior and development of this robot was almost identical to the robot in the first experiment as show in figures 7.3 and 7.4. For this reason we will not spend any time discussing this robot's early life but will rather move on to the second half of the experiment, when the behavior started to deviate.

Second Experiment: minutes 30-50

Much like the robot in the first experiment, at around 33 minutes into its development, this robot had adapted to its environment well enough to no longer need to spend the majority of its time looking for resources. The exception to this, shown in figures 7.3 and 7.4, occurs between the 40th-50th minute. Due to the increased interaction with objects as discussed shortly, the robot takes additional health damage as it learns how to properly interact, therefore it spends additional time during this period recovering it s health variable.

While the robot in the first environment spent much of its "free time" being idle simply due to a lack of anything to do i.e a very limited number of novelty sources to interact with, this robot had a much larger range of possible objects to learn about. As before, the robot's interest in the novel objects in the environment depended on the concentration of the nH_c and nH_s hormones. Initially, with a high value of the nH_c, the robot's attention was mostly focused on the randomly moving robots. During this period of high concentration of nH_c, in the initial instances the robot would simply engage in a following behavior moving behind the nearest moving robot. After around 2-3 instances of this following behavior, the robot began to intensify its interaction by engaging in both pushing as well as approaching the small robot from different angles. Since the randomly moving robots had been programmed to stop if contact was detected, the novelty value that the robot would associate with them greatly diminished over a period of around 5 minutes, dropping to almost zero novelty near the start of the 50th minute as shown in figure 7.5.

In contrast, with a medium concentration of the nH_c, the robot was attracted to the different arrangements of objects that were constructed with the small tin cans (see figure 7.1). Initially, the robot would either move close to these structures or slowly circle around them. After a couple of minutes, when the robot was familiar with the structures, it began to make physical contact with them through gentle bumps and pushes. Due to the lightweight nature of the tin cans, any physical contact from the robot would easily knock them over, and this resulted in the robot detecting not only a large amount of unexpected rapid movement around itself but also collisions, as some of the tin cans hit the robot. Since the robot only had a moderate amount of the nH_c when initially interacting with the structures, their falling resulted in

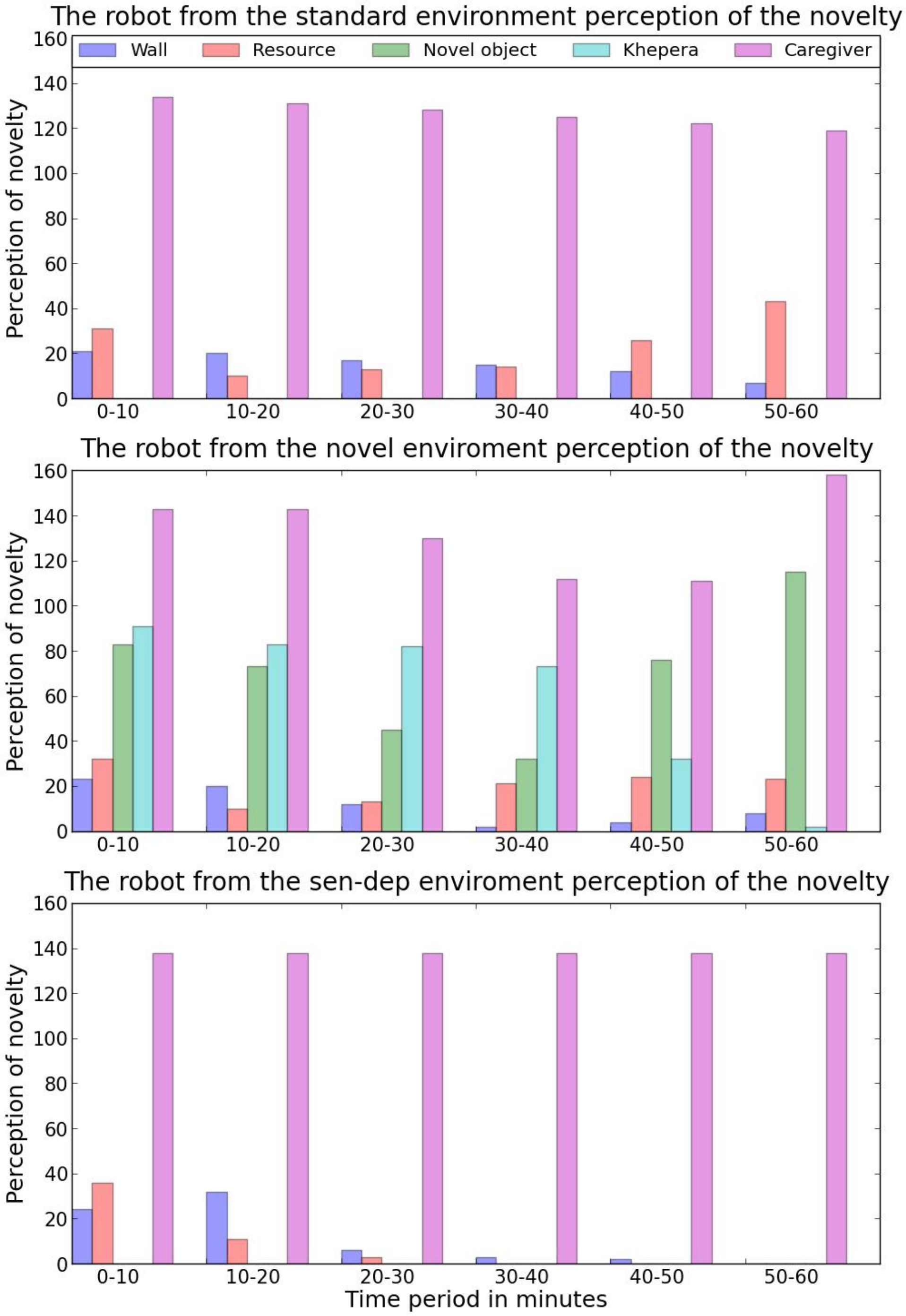


Fig. 7.5 The robots' perception of novelty
An overview of the average perceived novelty of 5 different aspects of the robot's environment during each time period. It should be expected that as a robot interacts with an aspect the novelty value will decrease. The exception to this is if the objects have unpredictable or dynamic behavior in which case the novelty value would be expected to rise as the robot interacts with it.

significant over-stimulation, leading to increased secretion of the nH_s and its withdrawal to a perceived safer location. These implications of the early contact with the structures resulted in the robot associating a higher level of perceived novelty with them due to the uncertainty of the outcome of any interaction. This increase in novelty associated with the structures along with the decrease in novelty associated with the kheperas resulted in structures having the highest perceived novelty as shown in figure 7.5. Due to the increased perceived novelty, the robot would now only interact with the novel structures with high nH_c levels. The higher concentration of nH_c protected the robot from becoming overstimulated due to unpredicted outcomes, which led to more thorough interaction with the structures. In the last 5 to 10 minutes of this period, the robot engaged with the structures in a number of different ways as it attempted to learn about them – including moving around them at different speeds, stopping near them at different distances, trying to move through them and pushing them with different intensities.

Second Experiment: minutes 50-60

At around 54 minutes into its development, the robot started displaying a new behavior: it would gently push over a structure before moving away and stopping. As we previously mentioned, when a structure was knocked down, the caregiver would replace it when the robot had moved away. As soon as the caregiver entered the environment to replace the tin cans, the robot immediately moved towards them and tried to interact with the caregiver. The caregiver, due to a number of factors such as size, shape and movement, was unsurprisingly perceived as highly novel by the robot (see figure 7.5). What was however interesting is the idea that the robot seemed to engage in this sequence of behavior “on purpose”. It is likely that, after experimenting with the objects, the robot has learned that by pushing the structures over, it could cause the caregiver to enter the environment and use this to satisfy its own desire for novelty. Before the 54th minute, the robot had not displayed this behavior sequence of trying to have the caregiver enter the environment; yet after the first occurrence, in the remaining 6 minutes of the experiment, this behavior occurred 11 additional times. In all cases, this behavior only occurred with high nH_c and low nH_s levels, supporting the idea that the robot is utilizing this behavioral sequence “on purpose” to satisfy its own desire for novelty. Examining the ENN seems to back up this idea as neurons associated with the caregiver are active when interacting with the tin cans.

This behavior by the robot could be regarded as the emergence of a form of tertiary circular reactions and potentially bear a similarity to a representation of cause and effect. With regard to tertiary circular reactions, the robot was demonstrating the ability to not only manipulate and experiment with different objects in its environment, but also to use these objects in order to change its environment, thus suggesting some sort of representation of cause and effect, an aspect of tertiary circular reactions [111].

The formation of these representations is less clear, though, since the robot's behavior of knocking over structures in order to bring the unseen caregiver back into the environment could potentially suggest object permanence, which we test later in section 7.2.3.

7.1.4 Third Experiment: Sensory deprivation

In the final experiment, instead of being allowed to roam in an open environment like the previous ones, after the first 10 minutes of interaction with the caregiver this robot was placed in a small cardboard box with the homeostatic resources directly in front of it, in an attempt to create a sensory deprivation experience (see figure 7.1). As would be expected, with both resources directly in front of it and little room to move, the robot remained mostly inactive throughout the sensory deprivation period as shown in figure 7.4.

7.2 Comparing the cognitive development of our robots

From the overview of the experiments, it appears that the robot that developed in the novel environment (Section 7.1.3) gained more advanced cognitive abilities than the robots developed in the standard and "sensory deprivation" environments. These advanced cognitive abilities would seem to support the idea that an environment which provides a richer sensorimotor experience over the course of development, leads to a greater development in cognition autonomous robots too. However, we must ask the question whether these more advanced cognitive abilities are a permanent result of the actual developmental process, or are simply an illusion or a temporal sensory phenomena due to the different environmental conditions. In order to try to understand if these developmental conditions had indeed affected the cognitive development of the robots, in this following section we will look to compare the robot under a range of different significant developmental events.

7.2.1 Comparison of neural development and activity

For a first comparison between the robots, we will look in closer detail at the development of their different neural networks, which can be seen in figure 7.6.

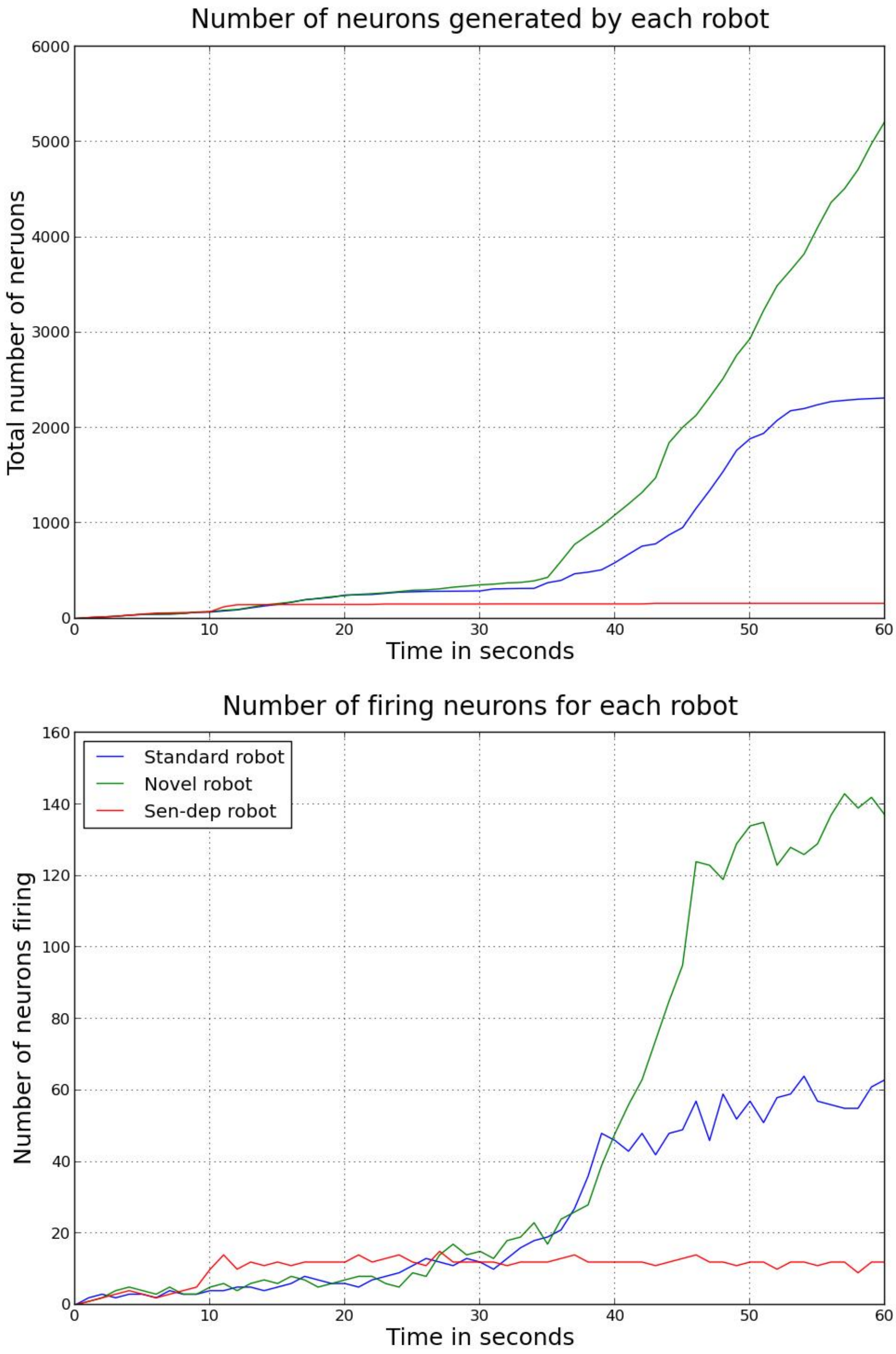


Fig. 7.6 The total number of neurons generated by the ENN for each of the robots over the course of the experiment (top) as well as the total number of firing nodes (bottom). The nodes themselves are generated as a function of the robots' interaction with its environment. A higher number of nodes would suggest that the robot has learned about a larger number of- or in more detail about different objects or aspects of its environment. A higher number of nodes firing appears to be related to the robot either noticing more aspects of the environment or having a greater understanding of the affordance of different aspects of the environment

Figure 7.6 shows the robot from the novel environment developed a larger neural network with significant growth occurring in the latter stages of the experiment, coinciding with the robot going through what we considered related to the coordination of secondary and tertiary reactions during the exploration period in section 7.1.3. Additionally, we can see again that the robot from the novel environment had a significantly larger number of neurons firing per action loop in the later stages. The increased number of nodes and neural activity from the novel robot can be explained due to this robot developing larger neural pathways. The increased pathways befitted the robot by allowing it greater understanding of its environment.

7.2.2 Learning and association: introduction of a new object

We next tested the ability of each of the three versions of the Koala robot – the robot from the experiments carried out in the "standard", "novel" and "sensory deprivation" environments – to learn by introducing two new novel objects that the robot had not seen before – two AIBO robots (one white and one black) shown in figure 7.1. These novel objects were set to work in a similar manner to the energy resource, recharging the energy of the robot when it was close, although these novel objects provided a much greater rate of energy replenishment – 30 units of energy a second, 4 times faster than the original energy resource. The Koala robot were then given a choice between the novel objects and the original energy resource, with the assumption that if/once the robots learned that the novel resources provided a greater charge, they would prefer it over the original resource.

In order to conduct this experiment, two minor changes were made to the robotic architecture. First, the energy level was set to 20% after every action loop, to ensure the robot had a permanent motivation to recover from energy deficits. Second, the secretion of the nH_c was suppressed to remove the motivation to move to the novel resources based purely on novelty value. The experiment involved two parts with results show in table 7.1 and figure 7.7.

For the first part of the experiment, the first novel object (the white AIBO) was placed directly in front of the Koala robot (close enough to charge) for a period of 10 seconds, to give the Koala an opportunity to learn about it; after this period, both this novel object and the original energy resource were placed slightly spread in front of the robot at a distance of around 1 meter, forcing the robot to choose which one to move to in order to replenish its energy levels. This entire cycle was then repeated 10 times.

Table 1. The choices[5]of the robots from the standard, novel and sensory-deprived environments between a novel and original resource in the first (left) and the second part (right) of the learning experiment discussed in section 7.2.2

First Experiment				Second Experiment			
Run	Standard Robot	Novel Robot	Sensory-dep Robot	Run	Standard Robot	Novel Robot	Sensory-dep Robot
1	Original	Novel	Original	1	Original	Novel	Original
2	Original	Novel	Original	2	Original	Novel	Original
3	Novel	Novel	Original	3	Original	Novel	Original
4	Novel	Novel	Original	4	Original	Novel	Original
5	Novel	Novel	Original	5	Original	Novel	Original
6	Novel	Novel	Original	6	Original	Novel	Original
7	Novel	Novel	Original	7	Original	Novel	Original
8	Original	Novel	Original	8	Original	Novel	Original
9	Novel	Novel	Original	9	Original	Novel	Original
10	Novel	Novel	Original	10	Original	Novel	Original

The results are reported in table 7.1, where we can see that the "novel" robot appears to immediately learn the increased energy affordance provided by the first novel object and was significantly more attracted to it. In comparison, the "standard" robot would often pick the novel resource after increased exposure to it, although as seen in figure 7.7 it was only slightly preferred. The "sensory deprived" robot did not show any signs of adaptation, systematically selecting the original energy resource.

For the second part of this experiment, conducted immediately after the first part, we changed the first novel object with the second (the black AIBO). Unlike in the previous part of the experiment, the new novel object was not placed in front of the robot at anytime, instead as before they were placed 1m ahead of the robot and slightly spread. Once again each of the versions of the Koala robot underwent another 10 runs with a same need to replenish its energy level. While the robot has never seen the second novel object before, it does possess similarities to the first, hence here we are testing if the robot can identify that the two novel objects share similarities and therefore may behave in a similar manner, i.e., both would offer rapid replenishment of the energy deficit. The results are shown in table 7.1, where we can see that even though the novel robot had never seen or interacted with the new novel object, unlike the other robots, due to its more developed neural network (Section 7.2.1), it is able to identify the similarities between the two novel objects and recognize that the second had similar properties (i.e., the ability to provide a rapid charge) to the first. In contrast while the

---

[5] Due to the closeness of the perceived energy affordance of the novel object and original energy resource for the standard robot (see figure 7.7), variation in perception (i.e the distance of the object) caused as a result of environmental noise may influence the perceived affordance and therefore the robots choice of 2 occasions (environmental noise can cause variation of perceived size of affordance by around 3%)

standard robot did seem to identify some similarities between the two novel objects, leading to a slight perceived affordance of energy recovery with the new novel objects, the perception was not great enough for it to choose the novel object over the safer original energy resource. Finally the sensory deprived robot which only showed minimal learning in the first stage of this experiment showed no association between the two objects.

7.2.3 Object Permanence: Recreation of a Hidden-Toy Test

One of the most commonly used methods for identifying and examining infant intelligence consists of the hidden toy test[6] [113, 10, 103] to see if the infant has gained a notion of object permanence. We reproduced this test by placing a needed resource in front of each of the Koala robots at a range of 2 meters. As the robot began to move towards the resource, 5 tins cans, used to build the previous novel structure shown in figure 7.1, were placed directly in front of the resource to block it from the robot's view. If the robot has a representation of object permanence, we would hypothesize that the robot would continue to move towards the object even when it is hidden from sight. If the robot stopped for more than 10 seconds, or 1 minute had passed after the resource had been hidden, the experiment ended to reduce the risk that the robot might find the resource accidentally or as part of its exploratory behavior. This experiment was conducted 10 times for each robot with the results shown in table 7.2.

As shown in table 7.2, the robots from the novel and standard environments both had some success in finding the resources once hidden; in comparison, the robot from the sensory deprivation environment was unsuccessful every time. If we look at the behaviour and neural activity of the robots, shown in figure 7.8, we can see that the robot from the novel environment was the only one to consistently search for the resource after it was hidden. In addition, this robot was also the only one which consistently (i.e., in every run) had activity along the neural pathway associated with the detection of the resource even after it had disappeared. This neural activity resulted from the fact that the original signal remained active along the pathway due to the modulation of this pathway by the different hormone concentrations, leading to feedback loops. These loops provide the robot with an ability akin to "active", or "short term" memory.

The 3 occasions when this robot failed to find the resource were due to the fact that the robot moved past the hidden resource. The fact that the neural activity remained high during these failed attempts suggests that, while the robot shows neural activity associated with the hidden

---

[6] We have also carried out experiments using the A-not-B test with robots that were developed in an environment with human "caregivers". Results of these experiments will be reported in a forthcoming publication.

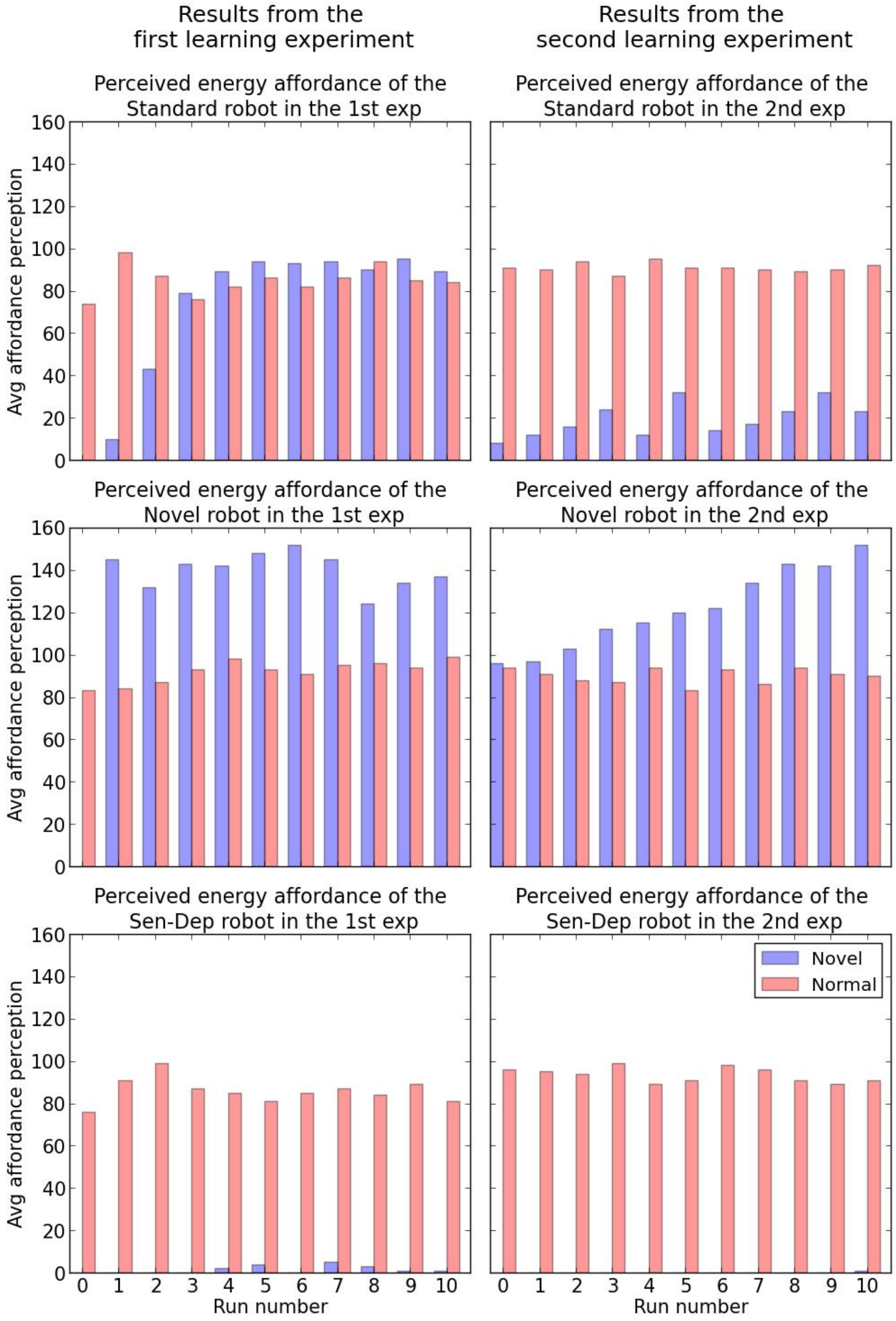


Fig. 7.7 The results of the first (left) and second (left) part of the learning experiment. The units in the y-axis show the strength of the perceived energy affordance of the two objects which is determined by equation 5.5. Run 0 is the perceived energy affordance before the start of experiments.

Table 2. Results of the Hidden-Toy Test of the robots from the standard, novel and sensory-deprived environments . Time is measured in seconds

Run	Novel Robot’s Time	Standard Robot’s Time	Sens-Dep Robot’s Time
1	32	16	Not found
2	14	Not found	Stopped
3	17	Not found	Stopped
4	32	32	Stopped
5	Not found	Not found	Stopped
6	Not Found	Not found	Stopped
7	15	Not found	Stopped
8	Not Found	43	Not found
9	23	Not found	Stopped
10	19	12	Stopped

object and the behavior it affords, without the expected feedback from sensory readings regarding the distance and position of the object, the robot cannot consistently locate it. This would appear to back up the previous observation that the first two robots had gained an ability consistent with the “understanding” of object permanence during their developmental runs, rather than having this skill from the start.

7.2.4 Violation of Expectation Paradigm

For the final experiment, we tested the robots using another common cognitive test, the Violation of Expectation paradigm (VOE). VOE experiments are normally carried out by showing very young infants two different pictures, one of which shows an impossible outcome – often some type optical illusion – while the other is almost identical but without the impossibility [127]. The experiment seeks to assess if the baby can notice the impossibility by measuring which picture it looks at more. The underlying assumption is that if the baby can identify the impossibility in the picture, it must possess some expectation about the object represented in that picture, and will look at it for longer than at the image without the impossibility.

We created a version of this experiment suitable for our robots. Here a white ball was placed in front of the robots. For the possible outcome we simply measured how long the ball which

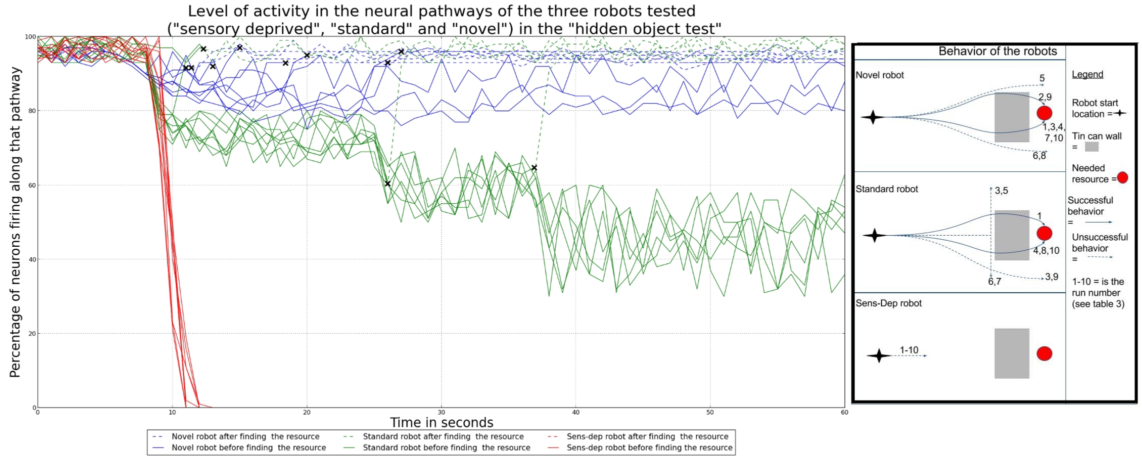


Figure 8. Overview of the behavior and neural activity of our three robots during the hidden toy test. The graph on the left shows the the neural activity along the pathways associated with the hidden resource of each robot. Neural activity is measured as the percentage of active nodes. Crosses indicate the points at which each robot has found (i.e., detected) the hidden resource. Dotted lines are used to show the neural activity of the robot after this point. In most cases, the robot will detect a resource some seconds before it physically reaches it. As we can see, the perception of the hidden resource gives rise to increased neural activity. After the resource is hidden (at around the 10-second mark), the pathway of the robot from the novel environment is the most active even without being able to see the resource: this robot has neural activity associated with the object and the behavior that it affords even after it has been hidden from view. On the right, the diagrams show the trajectories of the robots for each run. We can see that the robot from the novel environment was the most successful in finding the hidden resource.

has not been seen before held the robots attention. For the “VOE”, the white ball was again placed in front of the robot, however the robot’s sensors were modulated to make it appear as if the ball became smaller as the robot moved towards it, we once again measured how long the ball held the robots attention. If the robot can identify the “VOE” we would expect it to be hold its attention for a greater period of time.

As shown in table 7.3, the “VOE” held the novel robot’s attention for a significant duration in comparison to the possible object. In contrast, the robot from the standard environment only showed slightly more interest in the “VOE”, which due to the small difference it is not possible to conclusively suggest that this robot was showing an interest in the “VOE”. Finally, the sensory deprived robot showed no real difference in the time spent with both objects, suggesting the VOE paradigm had no real influence on the robot.

Our results suggest that the ability to respond to “VOE” arises as part of the later stages of the sensorimotor development process, which only the robot from the novel environment went

Table 3. Results from the VOE experiment showing the time (in seconds) that the robot from the novel environment, standard environment and sensory-deprived environment, spent focusing on or interacting with the impossible and possible object

Run	Novel Robot's Time		Standard Robot's Time		Sens-Dep Robot's Time	
	Impossible	Possible	Impossible	Possible	Impossible	Possible
1	46	17	21	18	19	20
2	48	15	22	19	23	21
3	43	17	19	22	21	23
4	49	18	17	15	20	20
5	54	14	19	12	15	15
6	52	18	16	15	19	20
7	49	19	15	14	18	19
8	56	16	14	13	19	20
9	49	14	16	16	14	13
10	54	17	21	15	18	16

through. Incidentally, this finding correlates with Piaget's developmental theory regarding when these skills should emerge. It should be noted that the VOE paradigm has been criticized by other developmental psychologists, who debate whether these skills are indeed learned as suggested by Piaget in his theory of development [113], or if they are part of a "core knowledge" possessed by all infants from birth, as suggested by Baillargeon [11, 10] and [129]. Both authors have previously used the VOE paradigm to demonstrate that babies are able to identify impossibility much earlier than would be expected if Piaget's theory was correct. In the case of a robot, we can be certain that this was not part of the robots "core knowledge" but is developed given the appropriate sensory-motor experiences.

There have also been debates (see [11, 10]) regarding whether the violation of expectation paradigm used with infants may have been biased suggesting that the impossible variation offered other additional stimuli (e.g. more activity or increased number of elements in the impossible picture), which attracts the infants' attention rather than their ability to identify or be attracted to the perceived impossibility. We could also imagine that the higher novelty of the impossible object might to some extent be responsible for the response of the infants. In the case of our robot experiments we would be happy to accept that the different response that the "novel" robot displays in the face of impossible experiences might be due to novelty. However, the novelty offered by the impossible object (a white ball which behaves in a way

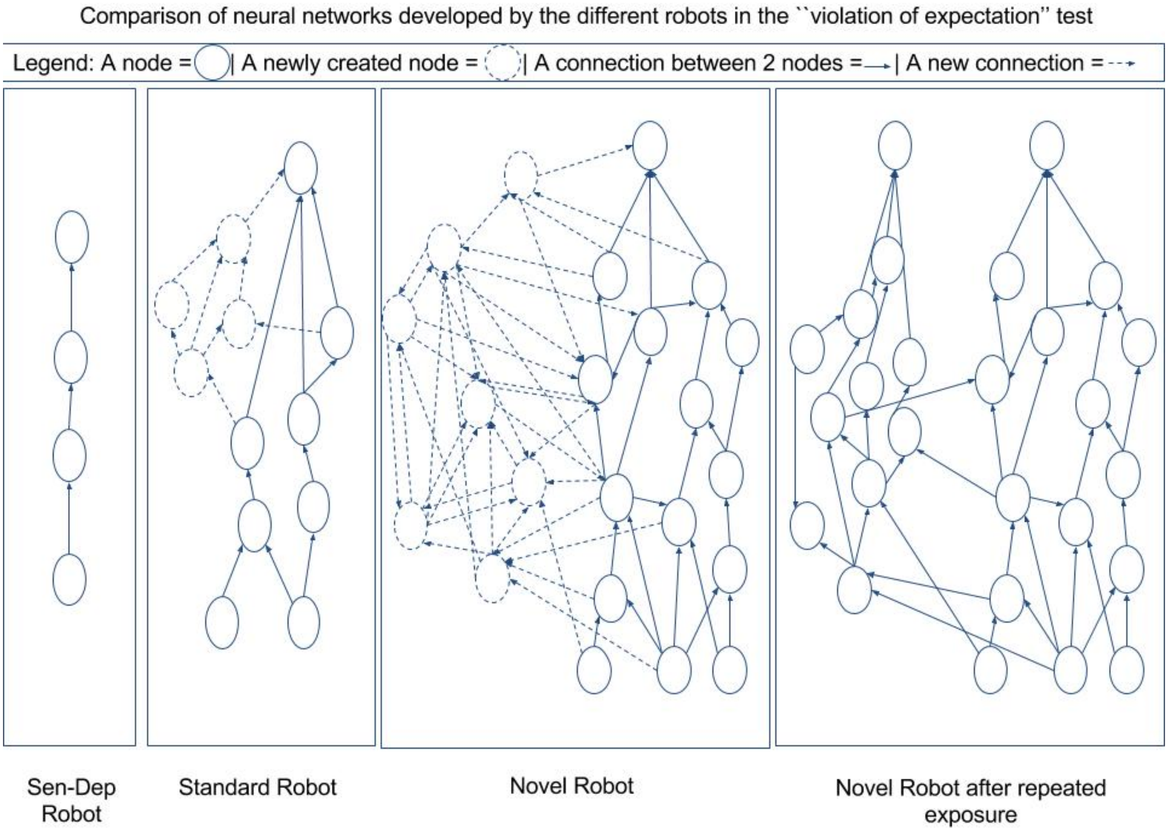


Figure 9.A snapshot of the neural pathway associated with a "violation of expectation" episode of the robots from the standard, novel and sensory-deprived environments . The greater effect can be seen in the network of the "Novel Robot": exposure leads to the generation of a larger number of nodes and connections leading to a greater neural activity and to the creation of a new pathway if exposure to the same type of episode re-occurs often enough.

that violates all the robot's previous sensory-motor experiences) is very different from the type of novelty offered by the perception of the novel object (white ball).

Although both exposure to a new object and a "VOE" episode produce "novelty", which we could define as the lack of behaviours and representations associated with an object, there is a qualitative difference in the effects that both experiences have on the robot's neural network. Novelty that arises from exposure to a new object is related to the level of plasticity between two connecting neurons (see equation 5.8): the higher the novelty of the object, the lower the level of plasticity, and a totally new object will give rise to new nodes and connections. Novelty related to a violation of expectation episode involves, in addition to the above, a change in the neural pathway associated with closely related previous experience. Specifically, an existing pathway is activated but a number of new nodes rapidly emerge along this pathway, linked to the elements of the new experience that violate the expectations from previous experiences, (see figure 7.9), label “Novel Robot”. Due to the activity of the pathway and the rapid growth of new nodes along it, a large number of “messy” and overlapping connections between nodes are quickly generated, increasing perceived novelty due to their high plasticity. These overlapping connections also effectively lead to the emergence of a sort of positive feedback loop within the pathway. This results in a greater level of activity along that pathway and therefore increased novelty. Over time, if the robot is repeatedly exposed to the “violation of expectation”, the original nodes and the new nodes will separate along two distinct pathways, the overlapping connections get “pruned”, and the feedback loops disappear (See section 5.2.3) (Figure 7.9, label “Novel Robot after repeated exposure”). This results in the consolidation of the new pathway and thus the previous “violation of expectation” becomes a normal “experience” for the robot. However, if exposure to the “VOE” stimulus is infrequent, the pathways will not split and the novelty associated with it will persist.

7.3 Conclusion

In this chapter we have investigated the role of sensorimotor experiences and environmental conditions in the emergence of more advanced cognitive abilities in the neuromodulated ENN. The robots endowed with this architecture showed increased “cognitive” abilities in relation to the quality of sensory motor experience that they were exposed to. Particularly, the robot raised in a context of sensory deprivation showed no additional behaviors or abilities outside simple reflex-like behavior that each robot started with. Further, this sensory deprived robot performed badly at adapting and in learning tasks compared to the other robots. Our model thus shows that a richer sensorimotor experience during early development correlates with greater “cognitive” ability in later life, as we hypothesised previously in chapter 6.

We have also shown how an autonomous robot implementing the architecture has the potential to go through developmental stages in a similar manner as outlined in Piaget’s sensory-motor theory. Our robot starts with a simple reflex-like behavior, yet through interactions with the external environment and stimulation from its internal environment, it develops more complex behaviors and cognitive abilities. With these new abilities built upon and expanding past behaviors.

Our robot was not explicitly designed around a developmental theory, but these developmental sub-stages emerged purely due to the interactions among the different aspects of the architecture – the hormonal, epigenetic, and ENN – and the external environment. In chapter3, when not all the previous components were present, the developmental phenomena mentioned here did not emerge. In particular, the addition of the ENN, with its learning representational capabilities, in interaction with the other elements of the architecture, was key factor in the emergence of these developmental phenomena. Our model thus offers potentially useful insights to bridge gaps between studies of epigenetic mechanisms such as [32] and developmental epigenetic theories such as [113] by showing how the former can lead to the emergence of the latter.

Chapter 8

Conclusion to the thesis

This chapter will give an overall summary on the research reported in this thesis. Section 8.1 provides a summary and comments on content of each chapter. Section 8.2 describes the key contributions of this research, while section 8.3.1 elucidates the future work.

8.1 Thesis Summary

In this thesis we have started to investigate the roles for potential epigenetic like mechanism in autonomous robots, with a focus on their adaptive benefits it terms of regulating developmental plasticity in regards to current environmental conditions. As we have discussed in chapter one, in the robotic community there is still a tendency to design task specific robotic controllers. These controllers which may well excel in a certain role within a specific environment are likely to struggle or even be rendered inoperable in the face of both environmental and internal change (see section 1.1).

In contrast, animals have been shown to adapt to changes in both their internal and external environments. While this is likely due to a combinations of a range of different mechanisms such as evolution, recent research has found evidence that an “Epigenetic’ mechanism, may play a potential critical role in the developmental process by regulating developmental plasticity. This potential has lead to a dramatic rise in interest into *“developmental epigenetics“* in both academia and general public interest.

New research into epigenetics (see section 1.2.2) has provided both a range of examples and potential mechanism in which different epigenetic mechanism can cause epigenetic changes in organism, effecting a range of different aspects of development. These changes have been shown to influence the emergence and developments in terms of cognition, behavior and morphology.

These epigenetic changes occurs as a result of the exposures to different internal and external stimuli. It is therefore possible that such changes may provide the organism with some type of adaptive benefit [1]. The ability to regulate the development through environmental exposures may therefore also provide befits in autonomous robots.

In chapter 2 we suggested a way in which a *"developmental epigenetic"* mechanism may be introduced into a robotic architecture. Our model uses the idea that hormones may provide signals which trigger and control the epigenetic process (see section 2.4.2). In this manuscript we have focused on one particularly type of epigenetic mechanism, the regulation of hormone receptors (see chapter 2). Namely up-regulation (increased sensitivity) and down-regulation (decreased sensitivity) on the receptors and how they can effect developmental plasticity. In order to test the epigenetic mechanism we also designed a homeostatic controlled robotic model. This controller saw the robot need to maintain 3 survival related variables of Energy, Health and Temperature within a predetermined boundary. Failure to do so would result in the robots death.

Our early work with such a hormone driven epigenetic mechanism and its effect on development variation is then reported in chapter 3. In these early experiment we tested the different robots in a range of environments. Here we found that the epigenetic mechanism led to significantly and consistently improved performance of a robot in comparison to other types of models. Not only were these robots better at maintaining their internal variables, but they were also able to develop a range of different behaviors in order to tackle the challenges they faced.

However, one issue we noticed during these experiments is the epigenetic robots still relied on the designer to provide information about the environment in order to operate. For example we needed to program the value of food or that hitting a wall would lead to a loss of health. In order to be truly adaptive we suggested autonomous robots must posses a form of open ended learning that allows them to continuously learn about their environments in a relevant way.

In order to achieve this we decided to test out a new type of neural network. We called this an emergent neural network or ENN. This network was designed to allow the robots to learn the "affordances" of different aspects of the environment. As we discussed in the implementation

---

[1] This is not suggest that all biological epigenetic mechanism are automatically beneficial. Particularly for modern humans in our rapidly changing environment, some of these mechanisms can clearly lead to undesirable effects. However, this is an issues best discussed in a different context.

of the ENN in chapter 5, this was achieved by having the robot learn to associate its own internal change with different sensory stimuli. For example, the robot was now able to learn that the blue ball provided energy by learning its energy levels increased when it was next to and facing the ball.

Due to the mechanism we implemented to control and modulate the emergence of different synaptic connections between the nodes the robot, over repeated interactions was able to learn which part of the environment caused the change. The choice to use the ENN rather than a another form of Network or learning mechanism was down to two main reasons

1. We wanted a learning system which was both open ended, and was flexible and free from different forms of tuning or need of parameters
2. We wanted a learning system which could easily be integrated into the epigenetic architecture, making a bespoke approach ideal.

As the ENN permits the robot to learn as a result of its interactions we needed to consider how best to generate learning situations. As we want the robot to be adaptive, it is important that its behaviors are appropriate given its current state and external environment. In order to achieve these desired behaviors, in chapter 4 we looked at extending our hormone model. Here we included two new hormones based on the roles of the steroid hormones Testosterone and Cortisol.

Research (see section 4.3) has found that in biological organism Testosterone and Cortisol are shown to regulate behaviors. Testosterone, is associated with outgoing behaviors. Cortisol, in associated with withdrawn behaviors. By implementing artificial variations through an artificial HPA (hypothalamic–pituitary–adrenal axis)-HPG (Hypothalamic–pituitary–gonadal) axis, we hope to create a novelty based behaviors system. In which our artificial testosterone hormone, which is secreted in relation to positive stimuli and high wellbeing, leads to outgoing and curiosity seeking behaviors. In contrast, our artificial Cortisol hormone, which in relation to low wellbeing and negative stimuli, leads to withdrawn behavior.

Experiments using this architecture showed that the mechanism had the desired results on behavioral modifications. A robot developed in a negative environment suppressed its curiosity-seeking behavior, preferring instead to simply stay in a safe location while maximizing its homeostatic levels, providing a buffer to help protect itself from it perceived environmental dangers.In contrast, a robot developed in a positive environment was more outgoing spending a greater portion of its time interacting with and exploring different sources of novelty.

The experiments conducted in chapter 4 showed the HPA-HPG mechanism seemingly providing the desired regulation of curiosity seeking behavior. Here we saw a robot exposed to significant levels of stress develop a more cautious approach to novelty and exploration. In contrast, a robot developed in a positive environment was more likely to interact with source of novelty. In addition, this "positively" developed robot was faster to recover from stressful experiences.

In chapter 6 we looked at how this system could be combined with the ENN to create our complete epigenetic learning model. Here we allowed different robots with the same architecture to develop in a range of different environments. The outcome of these experiments showed that each robot develop unique behaviors and characteristics. In addition, we demonstrated how each of these developed “phenotypes” were the most suitable for that robots environment. This would suggest that the mechanism is able to regulate developmental plasticity in a suitable manner given current environmental conditions.

In these experiments, for the first time we demonstrated how our architecture could allow a robot to enter a unknown environment with out any knowledge, yet learn and develop in a successful way. This included a robot in a hostile environment learning to not only avoid perceived aggressors, which could trick and beat them, but also developing behaviors such as playing dead. crucially this robot was also able to demonstrated that it was able to learn and adapt as its learned behaviors became less successful. This provided an early demonstration of the ongoing and open ended learning we desired

However, while we expected and hoped for this level of adaptation and development, we also noticed some interesting additional phenomena. This was the emergence of unplanned developmental phase. Each robot had at least two phases with the robots showing behavioral variation in each phases.These phases suggested to us that our model may not only allow the robot to develop unplanned behaviors, but also build upon them leading to the emergence of more advanced behaviors as the robot develops. Here we hypothesized that varying sensory-motor experience regulated these robots development.

The potential and existence of these developmental phases was tested in chapter 7. We investigated the roles of sensorimotor experiences and environmental conditions in the emergence of more advanced cognitive abilities. Here we once again placed our robot in a range of different environments with different levels of novelty. These variations were used in order to provide different levels of stimulation to each robot.

Immediately we noticed that the robot which was provided with greater stimuli seemed not only to develop faster but also showed signs of more advanced behavior, potentially even gaining certain cognitive abilities.

Interestingly, when looking into the potential developmental stages of the robot, we saw some similarities between its development and theories of human development. Specifically the robots seemed to go through a developmental process similar to one set out by Piaget when looking at infant behavior. Initially, like the infants, the robots behavior was limited to simple reflex behaviors. However, as the robot developed we saw the emergence of secondary and even tertiary circular behaviors.

As we saw behaviors associated with these different developmental stages we decided to examine these in more detail. We did this through the use of a range of tests used in infant cognitive assessment. According to Piaget[2], a child is expected to be able to complete the different assessments as they pass through different stages. Therefore we suggest that if the robot is able to complete a certain test it may have developed an analogy of a similar skill set. While we needed to make some minor adjustments to these test to account for significant differences between a robot and a child, we tried to keep the principles intact.

In these experiments we demonstrated that only the robot we identified as appearing to go through the different developmental stages, was able to pass all the tests. In contrast, the least developed robot which had been kept in a sensory-deprived state, was unable to pass any of the tests. This would suggest that the skills needed to pass the different test are gained by the robot as part of its developmental processes. Further all of these abilities were gained after an hour of run time. This may suggest that given longer periods, or more complex environments, it may be possible for the robots to develop further cognitive capabilities.

8.2 Contribution to Knowledge

Below we outline some of the contributions to knowledge that this work provides

- An epigenetic mechanism that uses hormones to regulates the robot's developmental plasticity based on its current internal and external state: Unlike many other forms of adaptation mechanism such as those described in section 2.1, the mechanisms proposed here can regulate the robot's behavior in relation to both internal and external changes. For example in chapters 6 & 7 we show how hormones can be used to regulate neural

---

[2] As we discuss in section 7.2.4, their has been debate over the correlation of these abilities and developmental stages

development in an appropriate manner.This is in contrast to many robotic models that are only able and designed to adapt to a changes in one or the other. For example a robotic controller may be designed to allow a robot to "eat" when its becomes hungry. However, these system rely on their being a constant food source for the robot to find. Simply put, if dynamics of the food source changes, i.e it becomes scarcer or harder to reach the robot cannot adapt to take into account this new variation. In comparison, our architecture in adaptive to both internal and external change. Meaning the robot not only knows to eat when it gets hungry, but it knows when to do so given its environmental conditions

- An open ended adaptation and learning system: The first, and main contribution to knowledge that this thesis has made is the introduction of an open-ended learning system. Here we demonstrated a potential model which has allowed real autonomous robots to adapt to a range of different dynamic and unstructured environments. This adaptation occurs without the need for any prior knowledge or tuning. Here, our model is able to learn the affordance of aspects of its environments as it develops and interacts with them. This also means that everything the robot learns is in relation to its own morphology and capabilities. This in comparison to a majority of robot controllers which are either highly task dependent or have limited automation.

- The need for survival: Again, unlike many other similar developmental models is the idea of survivability. Here we use three realistic survival variables which the robot must maintain in order to survive. Yet, while these are mostly simulated variables, they could easily be replaced with sensors to track the robot's actual state.

  In comparison, many other architecture do not consider and overlook the how shall we say, the "mortality" of the robot. Of course a robot is not simply going to die like a animal. However in a dynamic environments robots can become damaged, they can run out of power, if the robot overheats, circuits could melt. We therefore suggest that it is critical that all robotic architecture which aim for true automation must have system that incorporates the robots own survival. This is one cornerstones of our model, as we said in chapter 2, everything the robot does, it does to survive!

- Managing multiple goals and aims: Through our experiments, our model has demonstrated the ability to manage multiple goals. While initially this consisted of managing 3 survival variables, with the expansion of the model in chapter 3 and 4 the number of goals increased further. For example the inclusion of learning, curiosity and stress. Moreover, as we have stated previously, the robot can manage these goals in relation to its environmental conditions. Having an architecture able handle increased goals is relatively unique.

- Emergent behaviors and motivations: Another very unique aspects of our model is emergence of behaviors and motivations. Many similar models such as those discussed in chapter 1 and 2 use discreet behaviors and motivations (behaviors and motivations which are explicit programed for the robot to select). In contrast in our model we have neither, rather as discussed in chapters 2,6 and 7, here all behaviors and motivations emerge unplanned as the robot develops. The behaviors and motivations that emerge here could be considered fine grained, in the sense that behaviors and motivations occur as the result of different systems trying to modulate the robot.

- Emergence of developmental phenomena as the robot develops: Perhaps one of the most interesting contributions to knowledge that this thesis has made is the ability of our open-ended learning system to lead to the emergence of developmental stages and cognitive abilities. In this manuscript we have demonstrated a unique phenomenon where the robot is able to build upon it past experiences in order to becomes both more competent but also to develop new skills. While the research into understanding how this phenomenon has comes about is still in its relative infancy, we have seen highly promising results which we discussed in chapter 7.

    1. Since these capabilities are unplanned, we do not know how the robot may be able to advance given more time and stimuli. In comparison, robotic models such as those discussed in section 2.1 which are designed on a certain ability, may be able to replicate that ability, but are unlikely to progress further.

    2. Since the capabilities that we see here emerge unplanned and due to the combination and interaction of different components (such as the epigenetic mechanism and ENN), then perhaps our model may help understand how these abilities develop and arise in animals.

    3. Last, but not least, again due to the unplanned nature of these capabilities and the fact that they arise due to the interaction of the epigenetic system and other components, then perhaps this model may provide the start of a bridge between Piagetian epigenetics (and general developmental theories) and the recent new form of developmental epigenetics (see section 1.2.2. Here we have show how a plausible epigenetic mechanism can regulate developmental plasticity in a robot in away which leads to emergence of phenomena associated with Piagetian epigenetic ideas. This may potential indicate that these mechanisms may

- The effect of different neural structure related to the robots behavior: A final contribution to knowledge that we have made is the role of neural structures on behaviors and development. In chapter 7 we identified how the emergence of specific neural structures leads to significant changes in the robots behavior. For instance, in section 7.2 we showed how the emergence of hormonally modulated feedback loops in different neural pathways lead to the emergence of an “active memory” in the robot.

8.3 Current and Future work

The work presented here represents a portion of work carried into understanding and researching the potential and the effects of the proposed epigenetic learning system on an autonomous robot. Since the system has been developed,we have seen a range of potential avenues of investigation in regards to epigenetic modulated learning and development. Below we highlight a few of these avenues, some of which we have already begun looking into.

8.3.1 Further studies into the current epigenetic ENN

Firstly we consider some of the further work that should to be done with the current robotic model.

1. Further testing the cognitive ability of the epigenetic ENN:

   So far in this thesis we have demonstrated the importance of all components of the robotic model in the developmental process. We have also shown that the mechanisms leads to the emergence of developmental phenomena which bear some similarity to infant development. However, these finding came near the end of this research. So while we have found some very positive and interesting aspects to the model, their lies a large amount of work still to be done.

   Firstly, we have only just began looking at the cognitive capabilities of the robot. While it has so far demonstrated promising abilities there are both considerable more test and studies which should and need to be carried out to further understand the abilities of the robot. For example the A not B test (see [54]).

   Secondly, due to the complexity of the final neural network, new approaches need to be designed and implemented to help analysis it. For example only chapter 7 have we begun to identify how different neural structure play critical roles in the robots behavior. It is likely that there may be more of these structures that we have not yet identified that are significantly influencing the robots behavior.

2. Adaptation to morphology: In this thesis all our experiments have been undertaken using the Koala II robot (see figure 2.2). However in more recent work we have begun using other robots for varying reasons. Interestingly however, while the architecture needed some minor modifications to account for sensors and actuators the architecture was not only still functional but allowed for adaptation in regards to the robots specific morphology. Allowing for the robot to further exploit any ecological niche.

   We therefore suggest that this type of architecture will allow a robot to develop in a way that can exploit the advantages of its physical morphology. Therefore, it would interesting to study the effects of the architecture in different robots

3. Re-learning: In the manuscript our focus has been on how the architecture allows a robot to adapt and develop in regards to a certain environment. While we have placed the robots in new environments after being developed (see section 6), we have done so only as comparison to contrast behaviors between robots developed under different conditions.

   What we have not done however, is considered and looked into the potential for a developed architecture to adapt to a brand new environment. We suggest based on evidence from chapter 6 and 7 that as robots seems to build upon past experiences, these past experience will have some influence on the way the robot adapts to new environments.

   Due to the ever changing dynamics of real-world environments, if we want to push the epigenetic mechanism as a solution to real-world adaptation, it is important to examine and understand how the architecture can adapt to being exposed to new environments.

4. Human-robot interaction studies: The epigenetic system was not planed around human-robot interactions. However, as we have shown in chapter 6 the robot has performed and adapted well to different human interactions. It therefore is worth considering the potential applications the system bring to a human-robot interaction scenarios. For example, how certain care or household robots may use our system to adapt to the needs and behaviors of their user.

5. Platform to test developmental theories: Identifying and understanding development in humans and other animals is complicated, not least by the fact that we cannot simply open them up and look inside. As the epigenetic robot presented here has

some similarities to the developmental process as a infant, it could have potential applications to studying development. For example, we can look at how different experiences may influence development. For instance, both the robot and humans, when exposed extreme stress are shown to develop similar withdrawn phenotypes, which can result in both becoming essentially inoperable. This is not to say that the model provides a complete accurate representation but it may provide some insight and contributions.

8.3.2 Extension of the ENN

Secondly, we consider future work of how the ENN can be extended

1. Removal of classification algorithms: At the current time the ENN uses "Buckets" or pattern detection algorithms in order to categorize different stimuli (see section 5.2.2). For instance a color is determined to be red if a value is detected between a certain HSV range. However, we have actually found that the ENN can operate without these classification algorithms and can instead be feed raw data. Using this method the ENN is naturally able to categorize different stimulus through the emergence of different neural pathways. This means that the robot's perception and classification of a stimuli based upon its own experiences. Allowing both more categories to emerge (i.e multiple shades of reds each with a different meaning) and different boundary sizes.

   Generally this approach has led to some performance benefits, both in terms of raw survivability but also in the capabilities of the robot. For example using this later approach the robots is much better at learning and identifying different faces. However this system does have a problem of efficacy that needs to be addressed.

   - Firstly, under this approach many more nodes are created, leading to the network becoming too large and computationally intensive for a standard computer after around 2 hours run time. However, as the ENN has not been designed based on efficacy there is likely room for improvement in the implementation.
   - Secondly, and also contributing to efficacy issues is the problem of duplicate nodes, inefficient neural connections from rapid growth in addition to anomalies that occurring due to the pruning process (see section 5.2.2) As well as inefficiencies in computations, these issues can also lead to undesirable behaviors by the robot. However, this could lead to interesting of developmental disorder i.e., autism.

One approach we have begone investigating to alleviate this second problem is a sleep mechanism. Using this mechanism, the robot, when appropriate to do so will enter a state of sleep, effectively shutting down neural activity and leaving the robot with only basic reflex behaviors (see section 5.2.2) active. During this sleep process, a small system will analyze the network and look to repair, replace and remove poor, broken or unwanted nodes and neural pathways. So far this has shown to be highly effective. However, if for any reasons the robot needs to "wake" before the program is still running, then significant problems occur do unfinished modifications in pathways. This can range from simple issues such being unable to identify certain objects to losing the ability to move.

8.3.3 Group behavior and development

1. Indirect cross-generation transmission of epigenetic modifications: In our first chapters we suggested that in biological creatures epigenetic changes could be passed down through maternal care. In our own early investigations, we have seen some potential for a cross-generation transmission to occur in our epigenetic robots. In these early experiments we have placed robots which have been developed with exposures to certain stimuli, e.g extreme stress, in a neutral environments (i.e removal of the stressful environmental factor) along with new "undeveloped" robots. Here we noticed that through interactions between the robots, the developed robot's phenotype (i.e., withdrawn behavior if exposed to stress) would be passed on to the unexposed robots. So far we have noticed this phenomenon with both withdrawn and outgoing phenotypes.

   The question is however, what types of phenotypes can be passed on and what are the potential applications? In our mind we can see this providing an array of befits such as allowing robots to adapt to certain environments simply through interacting with other robots who have previously been exposed, giving these robots an advantage when they are later exposed to the same environment. Therefore, this phenomenon may provide an adaptive benefits to the autonomous robots

2. Epigenetic and social interaction regulation: In Lones *et al.* [81][3] we undertook a study looking at how groups of robots using the epigenetic mechanism would develop in a range of environments. One of the highlights of this work occurred in competitive environments. In these competitive environments we noticed the robots developing both advanced group behaviors and social structures. For example one group of robots

---

[3] the results of this study are not discussed in this manuscript

developed a “co-ordinated attack” against larger robots. In another groups we saw the emergence of social structure which included a dominant or "alpha" robot which would “lead” the other robots.

We believe that the epigenetic learning systems may facilitate more advanced group learning and adaptation than what we have seen here so far. From what we have seen in this study, in addition to the more advanced behaviors reported in chapters 6 and 7, we believe that the system proposed here may allow the groups of robots to develop into specialist roles give the current environment, group needs and the robots own morphology.

References

[1] (2001). *Introduction: Why epigenetic robotics*, volume 85.

[2] Angelov, P. (2012). *Autonomous learning systems: from data streams to knowledge in real-time*. John Wiley & Sons.

[3] Arkin, R. C. (1998). *Behavior-based robotics*. MIT press.

[4] Arnold, M. B. (1960). *Emotion and personality.* Columbia University Press.

[5] Ashby, W. (1952). Design for a brain. *Chapman and Hall.*

[6] Avila-García, O., Cañamero, L., and te Boekhorst, R. (2003). Analyzing the performance of "winner-take-all" and "voting-based" action selection policies within the two-resource problem. In *Alife*, volume 2801, pages 733–742.

[7] Avila-Garcia, O. and Cañamero, L. (2004). Using hormonal feedback to modulate action selection in a competitive scenario. *From animals to animats*, **8**, 243–252.

[8] Avila-Garcia, O. and Cañamero, L. (2005). Hormonal modulation of perception in motivation-based action selection architectures. In *Procs of the Symposium on Agents that Want and Like*. SSAISB.

[9] Bagot, R. C. and Meaney, M. J. (2010). Epigenetics and the biological basis of gene× environment interactions. *Journal of the American Academy of Child & Adolescent Psychiatry*, **49**(8), 752–771.

[10] Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. *Visual perception and cognition in infancy*, **23**, 265–315.

[11] Baillargeon, R., Spelke, E. S., and Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, **20**(3), 191–208.

[12] Bechara, A., Damasio, H., and Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, **10**(3), 295–307.

[13] Berthouze, L. and Metta, G. (2005). Epigenetic robotics: modelling cognitive development in robotic systems. *Cognitive Systems Research*, **6**(3), 189–192.

[14] Bird, A. P. (1985). CpG-rich islands and the function of DNA methylation. *Nature*, **321**(6067), 209–213.

[15] Braitenberg, V. (1984). *Vehicles: Explorations in synthetic psychology*. MIT Press.[MRWD].

[16] Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. MIT press.

[17] Breazeal, C. and Scassellati, B. (1999). How to build robots that make friends and influence people. In *Intelligent Robots and Systems, 1999. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference on*, volume 2, pages 858–863. IEEE.

[18] Breazeal, C. L. (2004). *Designing sociable robots*. MIT press.

[19] Buss, D. M. (2003). *The evolution of desire: Strategies of human mating*. Basic books.

[20] Cadet, J. L. (2016). Epigenetics of stress, addiction, and resilience: Therapeutic implications. *Molecular neurobiology*, **53**(1), 545–560.

[21] Cañamero, L. and Avila-García, O. (2007). A bottom-up investigation of emotional modulation in competitive scenarios. In *ACII*, pages 398–409.

[22] Cañamero, L., Avila-Garcia, O., and Hafner, E. (2002). First experiments relating behavior selection architectures to environmental complexity. In *Procs IEEE/RSJ Int Conf on Intelligent Robots and Systems*. IEEE.

[23] Cañamero, L. D. and Fredslund, J. (2000). How does it feel? emotional interaction with a humanoid lego robot. In *Proc. of American association for artificial intelligence fall symposium, FS-00-04*.

[24] Cangelosi, A. and Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots. *Cognitive science*, **30**(4), 673–689.

[25] Cangelosi, A., Schlesinger, M., and Smith, L. B. (2015). *Developmental robotics: From babies to robots*. MIT Press.

[26] Carvalho, J. T. and Nolfi, S. (2016). Behavioural plasticity in evolving robots. *Theory in Biosciences*, **135**(4), 201–216.

[27] Champagne, F. A. (2008). Epigenetic mechanisms and the transgenerational effects of maternal care. *Frontiers in neuroendocrinology*, **29**(3), 386–397.

[28] Champagne, F. A. and Rissman, E. F. (2011). Behavioral epigenetics: A new frontier in the study of hormones and behavior. *Hormones and behavior*, **59**(3), 277–278.

[29] Clemens, L. G., Gladue, B. A., and Coniglio, L. P. (1978). Prenatal endogenous androgenic influences on masculine sexual behavior and genital morphology in male and female rats. *Hormones and Behavior*, **10**(1), 40–53.

[30] Cortessis, V. K., Thomas, D. C., Levine, A. J., Breton, C. V., Mack, T. M., Siegmund, K. D., Haile, R. W., and Laird, P. W. (2012). Environmental epigenetics: prospects for studying epigenetic mediation of exposure–response relationships. *Human genetics*, **131**(10), 1565–1589.

[31] Crews, D. (2008). Epigenetics and its implications for behavioral neuroendocrinology. *Frontiers in neuroendocrinology*, **29**(3), 344–357.

[32] Crews, D. (2010). Epigenetics, brain, behavior, and the environment. *Hormones (Athens)*, **9**(1), 41–50.

[33] Crews, D. and McLachlan, J. A. (2006). Epigenetics, evolution, endocrine disruption, health, and disease. *Endocrinology*, **147**(6), s4–s10.

[34] Daitzman, R. and Zuckerman, M. (1980). Disinhibitory sensation seeking, personality and gonadal hormones. *Personality and Individual Differences*, **1**(2), 103–110.

[35] Darwin, C. (1962 (Originally published 1859)). *The Origin of Species by Means of Natural Selection Or the Preservation of Favoured Races in the Struggle for Life: With a New Foreword by George Gaylord Simpson*. Collier Books.

[36] Dauncey, M. (2014). Nutrition, the brain and cognitive decline: insights from epigenetics. *European journal of clinical nutrition*, **68**(11), 1179–1185.

[37] Day, J. J. and Sweatt, J. D. (2011). Epigenetic mechanisms in cognition. *Neuron*, **70**(5), 813–829.

[38] Del Giudice, M., Ellis, B. J., and Shirtcliff, E. A. (2011). The adaptive calibration model of stress responsivity. *Neuroscience & Biobehavioral Reviews*, **35**(7), 1562–1592.

[39] Der, R. and Martius, G. (2015). Novel plasticity rule can explain the development of sensorimotor intelligence. *Proceedings of the National Academy of Sciences*, **112**(45), E6224–E6232.

[40] Di Paolo, E. A. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. *Proc Simulation of Adaptive Behavior*.

[41] Di Paolo, E. A. (2003). Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. *Dynamical systems approach to embodiment and sociality*, pages 19–42.

[42] Dufty, A. M., Clobert, J., and Møller, A. P. (2002). Hormones, developmental plasticity and adaptation. *Trends in Ecology & Evolution*, **17**(4), 190–196.

[43] Fagiolini, M., Jensen, C. L., and Champagne, F. A. (2009). Epigenetic influences on brain development and plasticity. *Current opinion in neurobiology*, **19**(2), 207–212.

[44] Feil, R. and Fraga, M. F. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, **13**(2), 97–109.

[45] Finch, W. I. (1967). Geology of epigenetic uranium deposits in sandstone in the United states. Technical report, US Govt. Print. Off.,.

[46] Flores, K. B., Wolschin, F., and Amdam, G. V. (2013). The role of methylation of DNA in environmental adaptation. *Integrative and comparative biology*, page ict019.

[47] Fowden, A. L. and Forhead, A. J. (2009). Hormones as epigenetic signals in developmental programming. *Experimental Physiology*, **94**(6), 607–625.

[48] Fraga, M. F. *et al.* (2005a). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc natl Acad Sci USA*, **102**(30), 10604–10609.

[49] Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., *et al.* (2005b). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(30), 10604–10609.

[50] Francis, D., Diorio, J., Liu, D., and Meaney, M. J. (1999). Nongenomic transmission across generations of maternal behavior and stress responses in the rat. *Science*, **286**(5442), 1155–1158.

[51] French, R. L. and Cañamero, L. (2005). Introducing neuromodulation to a braitenberg vehicle. In *Procs 2005 IEEE Int Conf on Robotics and Automation*. IEEE.

[52] Frijda, N. (1986). *The Emotions*. Studies in Emotion and Social Interaction. Cambridge University Press.

[53] Gottlieb, G. (1991). Experiential canalization of behavioral development: theory. *Developmental psychology*, **27**(1), 4.

[54] Hackman, D. A. and Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in cognitive sciences*, **13**(2), 65–73.

[55] Hadorn, E. *et al.* (1961). *Developmental genetics and lethal factors.* London: Methuen & Co. Ltd.; New York: John Wiley & Sons, Inc.

[56] Haig, D. (2004). The (dual) origin of epigenetics. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 69, pages 67–70. Cold Spring Harbor Laboratory Press.

[57] Harper, L. (2005). Epigenetic inheritance and the intergenerational transfer of experience. *Psychological bulletin*, **131**(3), 340.

[58] Hiolle, A., Lewis, M., and Cañamero, L. ("2014"). "arousal regulation and affective adaptation to human responsiveness by a robot that explores and learns a novel environment". *"Frontiers in Neurorobotics"*, ("8").

[59] Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics*, **1**(2), 76–80.

[60] Holliday, R. and Pugh, J. E. (1975). DNA modification mechanisms and gene activity during development. *Science*, **187**(4173), 226–232.

[61] Hou, L., Zhang, X., Wang, D., and Baccarelli, A. (2011). Environmental chemical exposures and human epigenetics. *International journal of epidemiology*, page dyr154.

[62] Husbands, P., Smith, T., Jakobi, N., and O'Shea, M. (1998). Better living through chemistry: Evolving gasnets for robot control. *Connection Science*, **10**(3-4), 185–210.

[63] Insel, T. R. and Fernald, R. D. (2004). How the brain processes social information: Searching for the social **B**rain*. *Annu. Rev. Neurosci.*, **27**, 697–722.

[64] Jablonka, E. (2013). Behavioral epigenetics in ecological context. *Behavioral Ecology*, **24**(2), 325–326.

[65] Jablonka, E. and Lamb, M. J. (2002). The changing concept of epigenetics. *Annals of the New York Academy of Sciences*, **981**(1), 82–96.

[66] Jirtle, R. L. and Skinner, M. K. (2007). Environmental epigenomics and disease susceptibility. *Nature reviews genetics*, **8**(4), 253–262.

[67] Kaminsky, Z., Petronis, A., Wang, S.-C., Levine, B., Ghaffar, O., Floden, D., and Feinstein, A. (2008). Epigenetics of personality traits: an illustrative study of identical twins discordant for risk-taking behavior. *Twin Research and Human Genetics*, **11**(01), 1–11.

[68] Koolhaas, J. *et al.* (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews*, **23**(7), 925–935.

[69] Koolhaas, J., Van den Brink, T., Roozendal, B., and Boorsma, F. (1990). Medial amygdala and aggressive behavior. *Aggressive Behavior*, **16**, 223–229.

[70] Krichmar, J. L. (2012). Design principles for biologically inspired cognitive robotics. *Biologically Inspired Cognitive Architectures*, **1**, 73–81.

[71] Krichmar, J. L. (2013). A neurorobotic platform to test the influence of neuromodulatory signaling on anxious and curious behavior. *Frontiers in neurorobotics*, **7**.

[72] Krishnan, V., Han, M.-H., Graham, D. L., Berton, O., Renthal, W., Russo, S. J., LaPlant, Q., Graham, A., Lutter, M., Lagace, D. C., *et al.* (2007). Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell*, **131**(2), 391–404.

[73] Landgrave-Gómez, J., Mercado-Gómez, O., and Guevara-Guzmán, R. (2015). Epigenetic mechanisms in neurological and neurodegenerative diseases. *Frontiers in cellular neuroscience*, **9**.

[74] Ledón-Rettig, C. C., Richards, C. L., and Martin, L. B. (2013). Epigenetics for behavioral ecologists. *Behavioral Ecology*, **24**(2), 311–324.

[75] LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behavioural brain research*, **58**(1-2), 69–79.

[76] Lewin, K. (1931). *Environmental forces in child behavior and development.* Clark Univ. Press.

[77] Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., Sharma, S., Pearson, D., Plotsky, P. M., and Meaney, M. J. (1997). Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science*, **277**(5332), 1659–1662.

[78] Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological bulletin*, **116**(1), 75.

[79] Lones, J. and Cañamero, L. (2013). Epigenetic adaptation through hormone modulation in autonomous robots. In *Development and Learning and Epigenetic Robotics (ICDL), 2013 IEEE Third Joint International Conference on*, pages 1–6. IEEE.

[80] Lones, J., Lewis, M., and Cañamero, L. (2014a). Hormonal modulation of development and behaviour permits a robot to adapt to novel interactions. In *ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems*, volume 14, pages 184–191.

[81] Lones, J., Lewis, M., and Cañamero, L. (2014b). Hormonal modulation of interaction between autonomous agents. In *4th International Conference on Development and Learning and on Epigenetic Robotics*, pages 402–407. IEEE.

[82] Lorenz, K. (1935). Der Kumpan in der Umwelt des Vogels. *Journal of Ornithology*, **83**(3), 289–413.

[83] Lungarella, M., Metta, G., Pfeifer, R., and Sandini, G. (2003). Developmental robotics: a survey. *Connection Science*, **15**(4), 151–190.

[84] Macrì, S. and Würbel, H. (2006). Developmental plasticity of hpa and fear responses in rats: a critical review of the maternal mediation hypothesis. *Hormones and behavior*, **50**(5), 667–680.

[85] Malik, S., McGlone, F., Bedrossian, D., and Dagher, A. (2008). Ghrelin modulates brain activity in areas that control appetitive behavior. *Cell Metabolism*, **5**(0), 400–409.

[86] Maris, M. and Boekhorst, R. (1996). Exploiting physical constraints: Heap formation through behavioral error in a group of robots, iros'96, ieee. In *RSJ International Conference on Intelligent Robots and Systems*.

[87] Marques-Deak, A., Cizza, G., and Sternberg, E. (2005). Brain-immune interactions and disease susceptibility. *Molecular Psychiatry*, **10**(3), 239–250.

[88] Massot, M. and Aragón, P. (2013). Phenotypic resonance from a single meal in an insectivorous lizard. *Current Biology*, **23**(14), 1320–1323.

[89] Mazur, A. and Booth, A. (1998). Testosterone and dominance in men. *Behavioral and brain sciences*, **21**(03), 353–363.

[90] McEwen, B. S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences*, **840**(1), 33–44.

[91] McEwen, B. S. (2000). The neurobiology of stress: from serendipity to clinical relevance. *Brain research*, **886**(1), 172–189.

[92] McEwen, B. S. (2008). Understanding the potency of stressful early life experiences on brain and body function. *Metabolism*, **57**, S11–S15.

[93] McFarland, D. J. (1974). *Motivational control systems analysis*. Academic Press.

[94] McGee, S. L. and Hargreaves, M. (2011). Histone modifications and exercise adaptations. *Journal of applied physiology*, **110**(1), 258–263.

[95] McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonté, B., Szyf, M., Turecki, G., and Meaney, M. J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature neuroscience*, **12**(3), 342–348.

[96] McHale, G. and Husbands, P. (2004). Gasnets and other evolvable neural networks applied to bipedal locomotion. *From Animals to Animats*, **8**, 163–172.

[97] Mehta, P. H. and Beer, J. (2010). Neural mechanisms of the testosterone–aggression relation: The role of orbitofrontal cortex. *Journal of Cognitive Neuroscience*, **22**(10), 2357–2368.

[98] Meloni, M. and Testa, G. (2014). Scrutinizing the epigenetics revolution. *BioSocieties*, **9**(4), 431–456.

[99] Merrick, K. E. (2010). A comparative study of value systems for self-motivated exploration and learning by robots. *IEEE Transactions on Autonomous Mental Development*, **2**(2), 119–131.

[100] Miyakawa, H., Imai, M., Sugimoto, N., Ishikawa, Y., Ishikawa, A., Ishigaki, H., Okada, Y., Miyazaki, S., Koshikawa, S., Cornette, R., *et al.* (2010). Gene up-regulation in response to predator kairomones in the water flea, daphnia pulex. *BMC Developmental Biology*, **10**(1), 1.

[101] Moffitt, T. E. (2005). The new look of behavioral genetics in developmental psychopathology: gene-environment interplay in antisocial behaviors. *Psychological bulletin*, **131**(4), 533.

[102] Montoya, E. R., Terburg, D., Bos, P. A., and Van Honk, J. (2012). Testosterone, cortisol, and serotonin as key regulators of social aggression: A review and theoretical perspective. *Motivation and emotion*, **36**(1), 65–73.

[103] Munakata, Y. (2000). Challenges to the violation-of-expectation paradigm: Throwing the conceptual baby out with the perceptual processing bathwater? *Infancy*, **1**(4), 471–477.

[104] Nan, X., Ng, H.-H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998). Transcriptional repression by the methyl-cpg-binding protein mecp2 involves a histone deacetylase complex. *Nature*, **393**(6683), 386–389.

[105] Nijhout, H. F. (1999). Control mechanisms of polyphenic development in insects in polyphenic development, environmental factors alter some aspects of development in an orderly and predictable way. *Bioscience*, **49**(3), 181–192.

[106] Oberlander, T. *et al.* (2008a). Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (nr3c1) and infant cortisol stress responses. *Epigenetics*, **3**(2), 97–106.

[107] Oberlander, T. F., Weinberg, J., Papsdorf, M., Grunau, R., Misri, S., and Devlin, A. M. (2008b). Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (nr3c1) and infant cortisol stress responses. *Epigenetics*, **3**(2), 97–106.

[108] Oudeyer, P. (2015). What do we learn about development from baby robots? *CoRR*, **abs/1501.04796**.

[109] Oudeyer, P.-Y. and Smith, L. (2014). How evolution may work through curiosity-driven developmental process. *Topics Cogn. Sci.*

[110] Oudeyer, P.-Y., Kaplan, F., and Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, **11**(2), 265–286.

[111] Papalia, D. E., Olds, S. W., and Feldman, R. D. (1992). Human development edition.

[112] Pfeifer, R., Lungarella, M., and Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *science*, **318**(5853), 1088–1093.

[113] Piaget, J. (1952). The orgins of intelligence in children. *New york: international university press*.

[114] Piaget, J. (1978). Behavior and evolution (d. nicholson-smith, trans.) new york.

[115] Popma, A., Vermeiren, R., Geluk, C. A., Rinne, T., van den Brink, W., Knol, D. L., Jansen, L. M., Van Engeland, H., and Doreleijers, T. A. (2007). Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. *Biological psychiatry*, **61**(3), 405–411.

[116] Powledge, T. M. (2011). Behavioral epigenetics: How nurture shapes nature. *BioScience*, **61**(8), 588–592.

[117] Prince, C., Helder, N., and Hollich, G. (2005). *Ongoing emergence: A core concept in epigenetic robotics*. Lund University Cognitive Studies.

[118] R, H. (2006). Epigenetics: A historical overview. *Epigenetics*, **1**(2), 76–80.

[119] Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science*, **293**(5532), 1089–1093.

[120] Riggs, A. D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenetic and Genome Research*, **14**(1), 9–25.

[121] Russo, V. E., Martienssen, R. A., Riggs, A. D., *et al.* (1996). *Epigenetic mechanisms of gene regulation.* Cold Spring Harbor Laboratory Press.

[122] Schaffer, H. R. (1996). *Social development.* Blackwell Publishing.

[123] Shine, R. and Downes, S. J. (1999). Can pregnant lizards adjust their offspring phenotypes to environmental conditions? *Oecologia*, **119**(1), 1–8.

[124] Siegel, L. S. (1982). Reproductive, perinatal, and environmental factors as predictors of the cognitive and language development of preterm and full-term infants. *Child Development*, pages 963–973.

[125] Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of experimental psychology*, **81**(1), 174.

[126] Singh, S., Littman, M. L., Jong, N. K., Pardoe, D., and Stone, P. (2003). Learning predictive state representations. In *ICML*, pages 712–719.

[127] Sirois, S. and Mareschal, D. (2002). Models of habituation in infancy. *Trends in Cognitive Sciences*, **6**(7), 293–298.

[128] Song, S., Miller, K. D., and Abbott, L. F. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience*, **3**(9), 919–926.

[129] Spelke, E. S., Breinlinger, K., Macomber, J., and Jacobson, K. (1992). Origins of knowledge. *Psychological review*, **99**(4), 605.

[130] Stamps, J. and Krishnan, V. (2014). Individual differences in the potential and realized developmental plasticity of personality traits. *Frontiers in Ecology and Evolution*, **2**, 69.

[131] Stankiewicz, A. M., Swiergiel, A. H., and Lisowski, P. (2013). Epigenetics of stress adaptations in the brain. *Brain Research Bulletin*, **98**, 76–92.

[132] Steels, L. (1994a). The artificial life roots of artificial intelligence. *Artificial Life*, **1**(1-2), 76–110.

[133] Steels, L. (1994b). Emergent functionality in robotic agents through on-line evolution. In *artificial life IV: Proceedings of the fourth international workshop on the synthesis and simulation of living systems*, volume 4, page 8. MIT Press.

[134] Sterling, P. and Eyer, J. (1988). Allostasis: A new paradigm to explain arousal pathology.

[135] Sugano, S. and Shirai, Y. (2006). Robot design and environment design-waseda robot-house project. In *SICE-ICASE, 2006. International Joint Conference*, pages I–31. IEEE.

[136] Taylor, J. G. (1962). *The behavioral basis of perception.* Yale University Press.

[137] Thelen, E. and Smith, L. B. (1996). *A dynamic systems approach to the development of cognition and action.* MIT press.

[138] Thrun, S. and Mitchell, T. M. (1995). Lifelong robot learning. In *The biology and technology of intelligent autonomous agents*, pages 165–196. Springer.

[139] Tsankova, N., Renthal, W., Kumar, A., and Nestler, E. J. (2007). Epigenetic regulation in psychiatric disorders. *Nature Reviews Neuroscience*, **8**(5), 355–367.

[140] Tschacher, W. and Dauwalder, J.-P. (1999). *Dynamics, Synergetics, Autonomous Agents: Nonlinear Systems Approaches to Cognitive Psychology and Cognitive Science*, volume 8. World Scientific.

[141] Tyrrell, T. (1993). *Computational mechanisms for action selection*. Ph.D. thesis, University of Edinburgh Edinburgh, Scotland.

[142] Tzschentke, B. and Basta, D. (2002). Early development of neuronal hypothalamic thermosensitivity in birds: influence of epigenetic temperature adaptation. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **131**(4), 825–832.

[143] van Wingen, G., Mattern, C., Verkes, R. J., Buitelaar, J., and Fernández, G. (2010). Testosterone reduces amygdala–orbitofrontal cortex coupling. *Psychoneuroendocrinology*, **35**(1), 105–113.

[144] Waddington, C. (1936). Organizers in mammalian development. *Nature*, **138**(125).

[145] Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, **150**(3811), 563–565.

[146] West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

[147] Woolston, C. (2016). Researcher under fire for new yorker epigenetics article. *Nature*, **533**(7603), 295–295.

[148] Wyer, R. S. (2014). *The automaticity of everyday life: Advances in social cognition*, volume 10. Psychology Press.

[149] Zhang, X. and Ho, S.-M. (2011). Epigenetics meets endocrinology. *Journal of molecular endocrinology*, **46**(1), R11–R32.